POLITECNICO DI TORINO

I Facoltà di Ingegneria

Corso di Laurea Magistrale in Ingegneria Aerospaziale

Tesi di Laurea Magistrale

Approximations of Low-Thrust Trajectory Arcs by Means of Perturbative Approaches

Approssimazioni di traiettorie a bassa spinta mediante tecniche perturbative

Relatori:
Giulio AVANZINI
Mauro MASSARI

Candidato:
Alessandro PALMAS

ANNO ACCADEMICO 2009-2010

Ringraziamenti

Ringrazio il Dott. Massimiliano Vasile della University of Glasgow e il Dott. Federico Zuiani dello Space Advanced Research Team della University of Glasgow per il grande supporto nell'applicazione delle espansioni perturbative ai metodi di ottimizzazione diretta per l'ottimizzazione di traiettorie interplanetarie e per la disponibilità con la quale hanno condiviso con me i risultati di maggior interesse del lavoro.

Sommario

Il lavoro si propone di ottenere un'approssimazione analitica per archi di traiettoria a bassa spinta che permetta di rappresentare efficacemente lo scostamento rispetto all'orbita kepleriana standard. Sono state adottate due tecniche perturbative (i) singolare e (ii) standard, nelle quali il parametro di perturbazione, ε, è funzione dell'accelerazione fornita dal propulsore a bassa spinta, T/m e della accelerazione di gravità, μ/r^2. Il moto kepleriano fornisce i termini di ordine zero dell'espansione e i coefficienti per l'espansione al primo ordine sono funzione dei valori iniziali dei parametri orbitali e del modulo e direzione della spinta. Nel primo caso (perturbazione singolare) si opera direttamente sulle equazioni fondamentali della astrodinamica, l'approccio è semplice, ma limitato a problemi bidimensionali e fornisce soltanto la variazione secolare dei parametri orbitali. Il secondo approccio (perturbazione standard) opera sulle equazioni variazionali di Gauss, si trattano quindi problemi tridimensionali, identificando correttamente il drift degli elementi orbitali e la loro variazione sia di lungo che di corto periodo in una scala di tempo sub–orbitale.

Questa nuova tecnica è valutata in termini di accuratezza della soluzione e tempo computazionale comparando le sue performance con quelle ottenute da una procedura numerica standard (Metodo di Encke), e un'approssimazione analitica basata sulla trasformata di Fourier, dimostrantosi migliore degli altri in termini di tempo computazionale. L'espansione può essere usata per effettuare la propagazione orbitale per molte orbite, se il termine perturbativo rimane abbastanza piccolo, altrimenti gli archi di traiettoria devono essere ridotti e i coefficienti delle espansioni devono essere aggiornati per ridurre l'errore. L'uso dell'espansione per la discretizzazione di traiettorie interplanetarie è analizzato nel contesto dei metodi di ottimizzazione diretta per una ottimizzazione di traiettoria, si è riscontrato un considerevole risparmio di tempo computazionale per ottenere soluzioni ottimali non troppo accurate che rende il metodo interessante per ottenere velocemente soluzioni di primo tentativo per trasferte interplanetarie ottimali da fornire in ingresso a ottimizzatori più accurati.

Abstract

The objective of the work is to derive an analytical approximation for low–thrust trajectory arcs, that is, a mathematical representation of how the trajectory of a space vehicle deviates from a standard Keplerian orbit under the action of a low–thrust propulsion system. The objective is pursued by means of the application of both (i) singular and (ii) standard perturbation approaches, where the perturbation parameter, ε, is the ratio between the acceleration produced by the low–thrust engine, T/m, and gravity acceleration of the primary body, μ/r^2 and Keplerian motion provides the zero–order term of the expansion. Coefficients for the first–order perturbative terms are expressed as a function of initial values of orbit parameters and thrust intensity and direction. In the first case, the multiple scale method is adopted for directly expanding the solution to the fundamental equation of astrodynamics. The approach is simple, but it is limited to planar problems and it provides the secular variation of orbit parameters only. On the converse, the second approach, based on standard perturbation of Gauss's variational equations, allows for dealing with three–dimensional problems, correctly identifying both the drift in the orbital elements and their short–term variation on a sub–orbital time scale.

This novel approach is tested in terms of solution accuracy and computational time by comparing its performance with those obtained from a standard numerical procedure (e.g. Encke's integration method), and another analytical approximation based on Fourier transform. The perturbation method outperforms the others in terms of computational time, and the expansion can be used for propagating a low–thrust trajectory over several orbits, if the perturbing term remains sufficiently small. For higher values of ε, the trajectory arc must become suborbital and the coefficients of the expansion must be updated for keeping the error under a prescribed threshold. This motivates a careful analysis of the propagation error as a function of perturbing acceleration and trajectory arc amplitude, in order to identify the limits to the considered first order expansion. As a possible practical application of the standard perturbation approach, its use for the discretization of an interplanetary trajectory is considered in the framework of a trajectory optimization problem. The considerable savings in terms of computational time for deriving a fairly accurate optimal solution enhances the interest in the method, as a possible means for rapidly obtaining first–guess solutions for interplanetary optimal transfers, to be refined by more accurate propagation techniques once the most promising regions of the search–space have been identified by means of this faster propagation approach.

Contents

Approssimazioni di traiettorie a bassa spinta mediante tecniche perturbative

L'obbiettivo del lavoro è derivare un'approssimazione analitica per archi di traiettoria a bassa spinta, ossia una rappresentazione matematica della deviazione che subisce la traiettoria di un veicolo spaziale rispetto alla orbita kepleriana standard sotto l'azione di una propulsione a bassa spinta. L'obbiettivo è perseguito mediante l'applicazione di due tecniche perturbative (i) singolare e (ii) standard, nelle quali il parametro di perturbazione, ε, è dato dal rapporto tra l'accelerazione prodotta da un motore a bassa spinta, T/m, e la accelerazione di gravità causata dal corpo primario, μ/r^2 e il moto kepleriano fornisce i termini di ordine zero dell'espansione. I coefficienti per l'espansione al primo ordine sono espressi come funzione dei valori iniziali dei parametri orbitali e del modulo e direzione della spinta. Nel primo caso, la scala multipla è adottata per ottenere direttamente la soluzione delle equazioni fondamentali della astrodinamica. L'approccio è semplice, ma è limitato a problemi planari e fornisce soltanto la variazione secolare dei parametri orbitali. Per contro, il secondo approccio, basato sulla perturbazione standard delle equazioni variazionali di Gauss, permette di risolvere problemi tridimensionali, identificando correttamente il drift degli elementi orbitali e la loro variazione sia di lungo che di corto periodo in una scala di tempo sub–orbitale.

Questo approccio innovativo è testato in termini di accuratezza della soluzione e tempo computazionale comparando le sue performance con quelle ottenute da una procedura numerica standard (ad esempio il mediante il metodo di integrazione di Encke), e un'altra approssimazione analitica basata sulla trasformata di Fourier. Il metodo perturbativo si dimostra migliore degli altri in termini di tempo computazionale e l'espansione può essere usata per effettuare la propagazione orbitale per molte orbite, se il termine perturbativo rimane abbastanza piccolo. Per valori più alti di ε, gli archi di traiettoria

devono essere ridotti a distanze sub orbitali e i coefficienti delle espansioni devono essere aggiornati per tenere l'errore sotto una soglia prescritta. Questo motiva la attenta analisi dell'errore sulla propagazione orbitale come funzione del parametro perturbativo e dell'ampiezza dell'arco di traiettoria, in modo da identificare i limiti della espansione al primo ordine considerata in questa tesi. Come possibile applicazione pratica della perturbazione standard, il suo uso per per la discretizzazione di traiettorie interplanetarie è analizzato nel contesto dei metodi di ottimizzazione diretta per una ottimizzazione di traiettoria. Il considederevole risparmio di tempo computazionale per ottenere soluzioni ottimali non troppo accurate evidenzia gli aspetti interessanti del metodo, che potrebbe venire utilizzato come possibile strumento per ottenere velocemente soluzioni di primo tentativo per trasferte interplanetarie ottimali, da fornire in ingresso a ottimizzatori più accurati che possano calcolarle con maggiore precisione, una volta che le regioni più interessanti e promettenti dello spazio di ricerca sono state identificate per mezzo di questo veloce strumento basato sulle espansioni perturbative.

Caratteristiche Generali

Due differenti approssimazioni analitiche per traiettorie a bassa spinta sono state derivate per mezzo dell'applicazione di tecniche perturbative, entrambe usano come variabile indipendente una coordinata angolare. Il primo approccio introduce delle espansioni in funzione di θ nelle equazioni bidimensionali del moto mentre nel secondo le espansioni sono funzione della longitudine media L inserite nelle equazioni variazionali di Gauss. Il metodo della doppia scala è più semplice ma è limitato all'analisi di problemi piani mentre il secondo - perturbazioni standard - è più complesso ma tridimensionale.

I risultati ottenuti per mezzo delle espansioni perturbative nella valutazione della evoluzione dei parametri orbitali in traiettorie a bassa spinta sono soddisfacenti, specialmente se si considera il metodo di perturbazione standard, esso è capace di catturare i termini sia di lungo che di corto periodo e non è soggetto ad alcuna singolarità per orbite circolari e con inclinazione nulla, dal momento che la formulazione è basata su elementi equinoziali. La solidità delle formulazioni è stata provata per mezzo di molte simulazioni in cui i risultati venivano comparati con un'accurata soluzione numerica basata sul metodo di Encke. E' stato effetuato inoltre un altro confronto con una approssimazione analitica trovata in letteratura [18] basata sulle espansioni in serie di Fourier, che si è dimostrata migliore dal punto di vista del risparmio in tempo computazionale ma che coglie solo i termini di lungo periodo e necessità di un controllo periodico per poter applicare la trasformazione di Fourier.

Il risparmio in tempo computazionale rispetto alla soluzione completamente numerica è considerevole nel caso della propagazione orbitale, entrambe le formulazioni derivate consentono una riduzione sul tempo della CPU tra il 90 e il 98% se si vogliono calcolare tutti i punti dell'arco propagato, esattamente come fa un metodo numerico. Ma il metodo approssimato consente anche una valutazione diretta del punto finale di un dato arco, senza la necessità di calcolare tutti i punti intermedi, cosa che per un metodo numerico è imprescindibile.

Ci sono dei limiti che restringono la possibilità di applicare le espansioni trovate. In generale il rapporto ε tra l'accelerazione del veicolo e quella gravitazionale deve essere limitato affinchè si possano ottenere risultati più accurati oppure si deve ridurre la lunghezza massima dell'arco approssimato.

Riguardo le perturbazioni in doppia scala ci sono delle limitazioni specifiche:

- il modulo della spinta è assunto costante, così come la direzione della stessa è scelta o allineata con la velocità o fissa rispetto al sistema di riferimento radiale–trasversale;
- il rapporto tra l'accelerazione del veicolo e quella gravitazionale deve essere limitato affinchè si possano ottenere risultati più accurati oppure di deve ridurre la lunghezza massima dell'arco approssimato;
- la derivazione è stata ottenuta solamente per problemi bidimensionali;

Quest'ultimo aspetto è particolarmente rilevante dato che restringe considerevolmente il campo applicativo, impedendo di trattare missioni, o parti di esse, in cui sono presenti trasferimenti che implicano cambiamenti di piano.

Anche per la perturbazione standard il modulo della spinta è assunto costante e la sua direzione è fissa rispetto al sistema di riferimento radiale–trasversale–normale, ma per il resto la trattazione è del tutto generale.

Va sottolineato che le limiatazioni imposte dalla costanza del modulo della spinta e della sua direzione non rappresentano un limite eccessivo dal punto di vista applicativo dato che si può discretizzare la traiettoria e si possono assumono dei valori costanti per le variabili di controllo in ogni sotto–arco, la cui lunghezza sarà fissata in modo tale da limitare l'errore sulla propagazione dei parametri orbitali.

E' stata effettuata anche una accurata valutazione dell'errore, si è visto che la precisione delle espansioni è una funzione dell'accelerazione adimensionalizzata e ovviamente della lunghezza angolare dell'arco approssimato. Si è quindi analizzato il valore del parametro perturbativo ε in due contesti tipici, geocentrico ed eliocentrico, per una configurazione di veicolo moderno. Si è visto che con la corrente tecnologia in caso geocentrico i valori del

parametro perturbativo rimangono alquanto bassi e quindi possono venire approssimati archi anche abbastanza grandi; mentre nel caso di missioni interplanetarie lontane dal sole (ad esempio ad una distanza vicina a Giove) l'approssimazione può essere fatta per archi molto più piccoli a causa del valore elevato assunto dall'accelerazione adimensionalizzata. Va comunque detto che in caso di veicolo molto lontano dal sole il modello di spinta impulsiva può nuovamente essere considerato valido anche per la propulsione elettrica, dato che in tal caso il tempo per il quale viene acceso il motore diventa trascurabile così come la variazione di posizione, quindi l'approssimazione con spinta continua non è più necessaria.

Applicazione Preliminare ad un Problema di Ottimizzazione Diretta

Le approssimazioni analitiche trovate mediante la perturbazione standard sono state applicate ad un metodo di ottimizzazione diretta, il loro compito è stato sostanzialmente la sostituzione del propagatore d'orbita numerico, e potrebbero venir applicate anche per implementare il calcolo analitico delle derivate dei gradienti dell'indice di ottimizzazione e della Matrice Jaconiana, ma questo aspetto non viene preso in considerazione in questo studio. Sono state percorse due vie per l'implementazione delle espansioni, una con propagazione in avanti a partire dai punti iniziali e una centrata: è chiaro che per lo stesso cammino angolare totale l'espansione centrata propaghi per una minore distanza dal punto iniziale e quindi ci si aspetta una precisione migliore. I risultati confermano questo trend atteso, con l'approssimazione centrata si ottiene una precisione migliore di due o tre volte rispetto alla propagazione in avanti per la stessa lunghezza totale di arco approssimato.

Queste tecniche permettono un grande risparmio di tempo computazionale rispetto ad una integrazione puramente numerica e la loro principale applicazione potrebbe essere quella di implementare un veloce algoritmo capace di ottenere soluzioni ottime non molto accurate da fornire come soluzioni di primo tentativo ad un ottimizzatore numerico raffinato, avendo perlustrato lo spazio di ricerca della soluzione e identificato le zone di maggior interesse con il metodo approssimato.

Lavori Futuri e Estensioni

I passi futuri per la ricerca dell'applicazione di tali espansioni analitiche per la propagazione di orbite a bassa spinta possono essere descritti come segue:

- **Implementazione di un'efficace metodo di derivazione analitica**: la disponibilità si espansioni analitiche permette di ottenere delle espressioni analitiche per le derivate dell'indice di ottimizzazione e per la matrice Jacobiana.

- **Implementazione di una tecnica mista**: le tecniche doppia–scala e perturbazione standard possono essere integrate in un approccio misto che permetta di ottenere la stessa precisione usando sotto–archi più lunghi per via del minore drift legato al valore di ε di cui è affetta la doppia scala, che cattura solamente i termini di lungo periodo.

E' inoltre possibile pensare di ottenere una descrizione completamente tridimensionale per l'approccio doppia–scala in modo da allargare il suo campo applicativo a trasferimenti che descrivano un cambio piano; oppure si può cercare di ottenere una approssimazione mediante la perturbazione standard ma al second'ordine, in modo tale da aumentare la accuratezza e permettere l'utilizzo di archi più lunghi. Entrambi questi obbiettivi sembrano comunque alquanto complessi ma sarebbero capaci di garantire un notevole miglioramento per le potenziali applicazioni delle espansioni perturbative nel campo del progetto e analisi delle missioni spaziali.

Chapter 1

Introduction

1.1 Objective of the work

In this chapter are analyzed the logical steps which justify the work

- Interplanetary missions or mission with great energy changes or plane changes need a lot of fuel if conventional propulsion systems are used;
- The electric propulsion could represents a solution (high I_{sp}) but it has other kind of problems (low–thrust, high electric power requirements, power budget problems if coupled with solar panels);
- Low–thrust radically changes the trajectory nature: it is not a keplerian motion anymore with impulsive ΔV, but there are finite arcs along which the energy and/or the velocity are changed;
- This aspect complicates the difficult issue of optimizing the trajectory for a given mission objective;
- Analytical approximations could help:
  - to describe compactly the trajectories;
  - to facilitate the search of optimal solutions by means of the identification of good candidates;

Thus the aim of the work is to find some analytical approximations by means of which describing in a simple and compact way the trajectory of a low–thrust propelled vehicle.

1.2 Space Missions Design

The design phase in Space Missions is crucial for several reasons since it has the fundamental task of defining, all the major aspects of the mission. Given the complexity of the problem only a general outline of mission phases and related 'costs' will be initially derived, whereas a more accurate definition of every engineering aspect related to the mission will be refined during the design process. But in doing so, this phase is constrained by a lot of opposite needs, in fact if on one side the so called time–to–market require all the studies to be fast, the horizon of new futuristic objectives, requires a tight precision so a typical computational demanding process of accurate calculations.

In this framework, trajectory design represents one of the most challenging tasks because of the relevance of the analysis to be performed. It usually requires maximization or minimization of a performance index, for example maximize final mass or payload, minimize ΔV, that is propellant consumption, or trip time, for given mission requirements. The set of typical missions are

- transfer between specified orbits;
- rendezvous mission (transfer from a specified orbit to a definite time-dependent position on a target orbit);
- interplanetary missions (peculiar rendezvous missions);
- station keeping;

but often a complete mission shows more than one of the prevoius phases and more than one time each.

The problem of fuel consumption in particular is crucial when dealing with missions involving long expected operational times. The efficiency of the spacecraft acceleration can be measured by comparing the velocity increment ΔV with the characteristic velocity c of the propellant. Specific impulse I_{sp} is defined as

$$I_{sp} = \frac{c}{g_0} \tag{1.1}$$

where g_0 is the gravity acceleration measured on the Earth surface. From Tsiolkovsky equation [6] one can express

$$\frac{\Delta V}{g_0 I_{sp}} = \log \frac{m_1}{m_2} \tag{1.2}$$

where m_1 is the initial mass and m_2 the final one (after the engine burn). It is easy to see that as I_{sp} goes higher the mass reduction would decrease rapidly. If one considers a limit case as[1]

$$\Delta V \cong g_0 I_{sp} = c$$

the final mass would be

$$m_2 = \frac{m_1}{e} = 0.3679\ m_1$$

For a mission from Earth to Mars a ΔV as large as 18 km/s is required one way, so that in order to meet the requirement set by the previous equation a I_{sp} of 1800 s would be needed. The maximum value given by the most efficient chemical propulsion fuel-oxidizer combination (Hydrogen and oxigen) is about 450 s[2]. A different type of propulsion system is thus needed such as electrical propulsion. With this class of thrusters, I_{sp} from 1000 up to values as high as 5000 s can be reached[3]. Unfortunately, low–thrust propulsion systems are characterized also by some negative features:

- low values of thrust: high I_{sp} values are reached accelerating a very small amount of fuel; in general both the I_{sp} and propulsion system thermodynamic efficiency η_{el} are an inverse function of mass flow rate for given electrical power P_e, but since the thrust force is directly linked to this quantity

  $$T = \dot{m}_p c$$

  it is clear that this kind of engines can provide values of thrust that, in general, are small, typical values ranging from $10^{-5}\ N$ up to $10\ N$ maximum;

- power requirement: in order to have an acceptable efficiency η_{el}, high values of current and/or voltage are needed; this means in turn that the power requirement could easily be as high as few MW. It could be fullfilled by using the system in the quasi-stationary regime with the aid of accumulators, but even with a Δt_{on} equal to the 10% of the duty cycle, hunderd of kilowatts are needed. This requirement could not

[1] This is a typical example used to show how the ratio $\Delta V/c$ is important for having a sufficient final mass.

[2] The most efficient combination is given by Hydrogen and Fluorure but it is not used because of handling and toxicity problems.

[3] There are some unprototyped projects which can reach I_{sp} even higher of about $10^4\ s$ but they are not ready yet and the thrust values are so small that cannot be adopted for interplanetary missions.

be satisfied by means of solar panels, especially for mission far away from the Sun so the common idea is to move towards nuclear power generator[4];

- critical operative conditions: I_{sp} and efficiency η_{el} usually are also a function of fuel temperature in the electro-thermal thrust but even if they are function of the plasma current this cause a temperature rise due to the Joule effect and since the materials could stand to maximum steady temperature of about 2000 K this constraint results into another limit to engine performance. Also in this case the quasi-stationary regime could help but it does not eliminate the problem;

- unqualified and experimental equipment: finally the non-existence of qualified systems is also an important limit of this kind of propulsion, in terms of system reliability

Beside the issues related to system hardware and overall vehicle architecture, which are significantly affected by the use of this class of thrusters instead of a more conventional system based on chemical rockets engines, the need of lower consumption make necessary an effective way for analyzing all the mission trajectory manouvers. It can be done in different ways and by means of different methods (direct and indirect ones, they are described later on): optimization could mean

- constrained optimization: identify a performance index (fuel consumption or trip time to be minimized for example), fix some primary contrains (e.g. maximum trip time or maximum fuel consumption respectively) and secondary ones (maximum g-load in a given direction, maximum total ionizing dose absorbed, etc.) and look for a, possibly global, optimal solution;

- constrained multi-objective optimization: it is very similar to the previous one but more than one performance index is considered at the same time and this make the whole process more difficult but it should also be a more effective way of finding real optimal solutions.

These ways could be faced by means of different approaches, without entering in details here, there are three big families of optimization methods, indirect, gradient based (they are colled direct in what follows) and evolutionary algorithms, each one has different pros and cons, but all of them are widely adopted.

[4]In this case environmental and and safety issues must be faced.

There are a lot of examples of this kind of research in literature, in particular for the so called direct optimization field with the use of the DFET technique (Direct Transcription by Finite Element in Time) the works of Vasile and Bernelli Zazzera (cf. [21], [22], [23] and [24]) are a typical reference point, in their works various mission are used as a benchmark, usually they assume a thrust with a variable orientation and magnitude which is assumed to be dependent on the Sun distance (variable electric power income) and consider also fly-bys manouvers. The approach is the typical one encountered in direct methods[5], the states and the controls are discretized along the whole trajectory and the final profile is obtained by means of interpolation.

For the indirect optimization there are also a lot of examples in which the electrical propulsion is adopted, in particular here it has been reported the work of Casalino (cf. [5]), also in these case fly–bys are taken into account, but this time the states and controls are directly derived as continue functions.

The optimization for low–thrust trajectories by means of evolutionary methods could not be performed due to the number of variable too high which comes out from the discretization of the trajectory that is incompatible with those methods.

The optimization process is composed , for any type of optimization method, by some analytical relations but also many numerical evaluations so their computational time is strictly dependent on the progress of the computers velocity; the whole process would become faster and faster if some part of the numerical calculation could be performed analitically, or at least with some very good analitical approximations. In addition there are some other difficulties related to interplanetary trajectories optimization related to the research domain so particular strategies are needed:these are the main reasons of the work aim.

1.3 Elements of the Analysis

1.3.1 Basic Space Flight Mechanics

The motion of a spacecraft of mass m under the influence of a big mass $M \gg m$ (either a planet or the Sun) is known to be a conic section, from Newton classical mechanics and under a set of simplifying assumptions. The simplest approach to the description of the resulting motion neglects all other forces, like gravitational influence from other bodies, solar radiation pressure,

[5]The reason of this will be clear after the dedicated section.

aerodynamic drag, and so on; resulting in the so called restricted[6] two body problem. The resulting equation of motion is

$$\ddot{\vec{\boldsymbol{r}}} = -\frac{\mu}{r^3}\vec{\boldsymbol{r}}$$

where $\vec{\boldsymbol{r}}$ is the distance vector between the spacecraft and the main body and $\mu = GM$ is the gravitational parameter.

Once initial conditions (e.g. position and velocity at time t_0) are known, the trajectory of the vehicle can be described by an analitical relation, for the motion in the (fixed) orbital plane. A three dimensional trajectory is obtained once the position of the orbital plane with respect to a fixed inertial frame is defined. A total of six constants, the so called orbital parameters, provide a complete geometrical description of the spacecraft motion in the absence of perturbative forces as a funtion of the six components of initial position and velocity vectors[7]. Given these quantities one can evaluate forward or backward in time spacecraft position with respect the assumed intertial frame, the only numerical procedure being necessary for inverting Kepler's time equation in those cases when orbit propagation at a given instant is considered.

During a space mission it is ofted ncessary to transfer the spacecraft from an initial orbit to another one, e.g. for performing an orbit menoeuvre for reaching the final operational orbit from a transfer orbit; for injecting an interplanetary probe onto an hyperbolic orbit for leaving the Earth's sphere of influence; or just for compensating the effects of perturbations that slowly affects the orbit flown by the spacecraft, making it drift awawy form the desired operational one.

In all these cases it is necessary to change the velocity vector, its magnitude and/or its direction. The magnitude of velocity variation ΔV is directly related by means of Tsiolkovsky's equation with fuel consumption, so that minimizing the total ΔV for performing a given mission or a mission task, will result in a beneficial reductino of the fuel required, which in turn means either (i) a higher payload weight for a given launcher size or a(ii) a smaller overall vehicle weight at launch which may allow for saving on launch cost, or (iii) a longer operational life for the vehicle

If the ΔV is applied by means of chemical propulsion an impulsive approximation is usually valid, where the ΔT_{on} during which the thruster is turned on is assumed to be a negligible fraction of the orbit period, and the spacecraft position can be considered as approximately constant. The

[6]Remember that the first assumption $m \ll M$.

[7]The dynamic system in composed by a second order vector differential equation, so that six initial conditions are needed to have a unique solution.

velocity thus changes almost instantaniously and the knowledge of direction and magnitude of the ΔV is sufficient for easily deriving the new trajectory by means of analitical determination of the new orbit parameters of the trajectory that results from the application of the considered ΔV. In this framework it is clear how, when dealing with chemical propulsion most of the relations for the spacecraft motion in space are analytical.

Things become significantly more complex when electric propulsion is used, in which case the typical ΔT_{on} can become of the same order of magnitude as the orbital period, and the position can no longer be considered as constant. Low thrust could be considered as a perturbation acting on the spacecraft, slowly changing its trajectory.

In order to evaluate how the vehicle would move under the influence of such a force one must integrate numerically the equation of motion for the perturbed case

$$\ddot{\vec{\boldsymbol{r}}} = -\frac{\mu}{r^3}\vec{\boldsymbol{r}} + \vec{\boldsymbol{a}}_p \tag{1.3}$$

In many applications the perturbing term is due to the presence of the mentioned simplifying assumtions that neglects less significant forces, such that $\vec{\boldsymbol{a}_p} \ll \mu/r^2$ where $\vec{\boldsymbol{a}}_p$ is the perturbative acceleration. If the acceleration obtained from the electrical propuslion system is small compared to gravitational one it can be trated as a general perturbation acting on the satellite and all the methods derived in this field can be applied in a straightforward way.

When a perturbation is small compared with gravity acceleration, the resulting numerical problem is not simple to integrate if written in the previous form. There are many approaches to the problem of orbit propagation in the presence of perturbations:

- Cowell's Method
- Encke's Method
- Gauss's Variational Equations

The first one consist simply in writing the equation of motion in spherical coordinates. In this way all the three variables have a comparable rate of change and the numerical integration can procede well provided that a careful choice of integration step and order of the numerical scheme are selected.

The second method splits the problem into two parts: in the first one only the unperturbed trajectory in considered, i.e. the analitical evolution of the radius in time is known since it is a Keplerian motion; and the second part consist in integrating in time the difference $\delta\vec{\boldsymbol{r}}$ between the unperturbed

radius and the actual one in the presence of perturbations, which has its own expression for the dynamic equation and is suitable for numerical integration after some wise manipulations (see Appendix A for details).

The third method is based on a different point of view: the dynamic system is transformed from the $\vec{r}$ and $\vec{v}$ equations to a set of six equations for the orbital parameters. These quantities would remain constant in the of absence of perturbations whereas in case of disturbance they slowly vary and a numerical integration scheme is effective in estimating this variation as a function of perturbing force expressed in terms of radial, transverse and normal components.

At this point it is clear that, in case a low thrust propultion system is used, the simple evaluation of the trajectory requires a numerical integration which is usually a computational demanding process. If the orbit propagation scheme is integrated in an optimization method in which the trajectory evaluation is performed 10^4 times, in order to investigate optimal control force profiles for mission objectives, computational cost and time can become hardy feasible.

It is thus obvious that the derivation ot an analytical (although approximated) expression for the evolution of orbit parameters under the action of a low–thrust propultion system is of great interest for both the representation of the resulting motino in terms of a limited number of relevant parameters, and for obtaining fast and computationally efficient orbit propatgation tools to be used for preliminary mission design.

1.3.2 Quick Introduction to Perturbation Methods

There are a lot of examples of systems in which the mathematical model present a relation not analitically solveble due to the presence of a very small term which, if eliminated, allow to reach an analitical closed form solution. Appendix B presents few systems of this kind, that are classical examples such as oscillators with small damping or small masses. In more general terms, the complex procedure of perturbation methods could be outlined as follows: the system solution is supposed to be dependent on the indipendent variable and on the small parameter ε, $x = \phi(\varepsilon)$. One thus expands the solution in terms of a function of ε, usually a power series of ε,

$$\phi(\varepsilon) = \phi_0 + \varepsilon\phi_1 + \varepsilon^2\phi_2 + ...$$

and substitute this expansion into the equation of the system of differential equations.

A new set of equations is derived imposing equality for terms of the same order of magnitude, by which determining the quantities ϕ_i in order to have

a consistent solution. Note that the knowledge ot the solution $x_0 = \phi_0$ for $\varepsilon = 0$ is postulated and necessary for starting the precedure.

In order to make this procedure more clear an example presented in more details in Appendix B on Kepler's time equation is here , based on application of the perturbative expansions to an algebraic relation. The equation is

$$E - e \sin E = tn. \tag{1.4}$$

It is easy to see that if the eccentricity is zero the solution is trivial, so one can think to look for a solution $E(t; e)$ for $e \to 0$, which can be useful for near circular orbits since in these cases it is $0 \leq e \ll 1$. The solution is assumed to be given in the form

$$E(t; e) = E_0(t; 0) + \phi_2(e) E_2(t) + \phi_3(e) E_3(t) + o(\phi_3(e))$$

and the dependence on the small parameter e is exploited by the functions $\phi_k = e^{k-1}$, so that, for a second order accurate expansion, one gets

$$E(t; e) = E_0(t; 0) + eE_2(t) + e^2 E_3(t) + o\left(e^2\right)$$

Substituting it into equation (1.4) the equation achieves the form

$$\left(E_0 + eE_2 + e^2 E_3\right) - e \sin\left(E_0 + eE_2 + e^2 E_3\right) = tn$$

where introducing the McLaurin series for the sine function and collecting the different quantities one obtains with the same order for e

$$E_0 + e(E_2 - \sin E_0) + e^2 (E_3 - E_2 \cos E_0) = nt$$

The previous mentioned tuning relations are thus obtained by comparing the terms in the two sides of the equation, that is

$$E_0 = nt, \quad E_2 = \sin E_0, \quad E_3 = E_2 \cos E_0, \ ...$$

The solution found thanks to this expansion is

$$E(t; e) = nt + e \sin(nt) + e^2 \sin(nt) \cos(nt)$$

in figure 1.1 is plotted the erro between the exact solution and the approximated one.

This example it is effective for making clear how the perturbative expansion works. Togheter with the calculation, a good understanding of the limits and convergence issues must be considered since the formal process

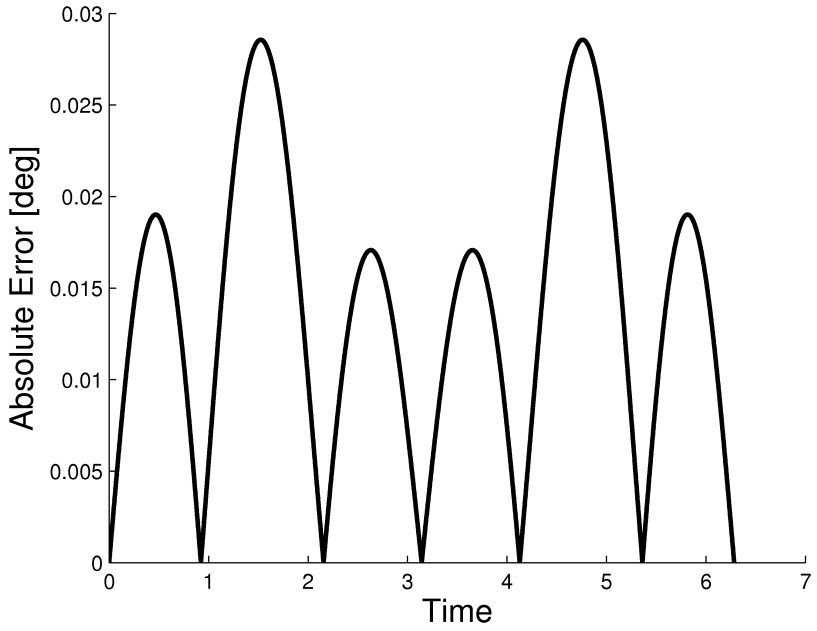

Figure 1.1: Error between numerical (exact) solution and approximated one

must be consistently applied with all its implications also when not intuitive equations are obtained.

In Appendix B all the topics concerning convergence, validity of the expansions, etc. are treated accurately. What is important to underline here is that, inside the boundaries which assure the correctness of the procedure from the mathematical standpoint the results can be considered consistent and this is what is assumed in what follows.

The same kind of process seen before can be applied to both ordinary and partial differential equations, as the weakly nonlinear oscillator presented in Appendix B. In case of problems in which the equation is an oscillator sometimes a particular procedure must be applied, the so called multiple scale approach (see again Appendix B), by means of which the behavior of armonic solution could be captured. In these cases the common procedure would fail since the limit for periodic solutions does not exist in genral.

This perturbative approach will be used in the sequel for low–thrust orbit propagation problems, since the perturbative acceleration can be considered a small parameter, a zero–order solution for ε is available, i.e. Keplerian motion, and the availability of an approximated solution dependent on the perturbative parameter can be hypothesized.

1.3.3 Optimization Algorithms in Space Mission Analysis and Design

Finding an optimal solution is surely one of the major engineering goals: once the objectives are chosen the solution being looked for is the one which allow to maximize the performance index which is a quantity that measures how well the solution reaches the imposed objectives.

There are a lot of optimization algorithms that has been developed specifically for application to space mission analysis, the most important features of which are reported in Appendix C with some technical discussions of the main aspects. A brief outlook focused on the most relevant characteristic for the objectives of the present work is outlined in this paragraph.

There are three families of optimization algorithms:

- Indirect Optimization Methods;
- Gradient–Based Optimization Methods;
- Evolutionary Methods.

Indirect Optimization Methods: when continous problems are dealth with, optimality conditions can be derived once the boundaries and the dynamic model are assigned. These conditions are derived forcing to zero the first order variation of the performance index, as it is done with the first derivative i.e. stationary condition, but also a second variation condition should be considered. In this way the problem does not need to be discretized and the optimal solution is found without evaluating the performance index directly, for this reason the name indirect. The problems are the dependence of the optimal solution on the initial guess and the possibility of finding sub-optimal solutions.

Gradient–Based Optimization Methods: continous problems are discretized, a multiple shooting problem is formalized and the optimal solution is found after a first guess (so also in this case the result is affected by this choice) moving towards the direction in which the performance index is maximized, this direction is found computing, by means of the perturbative approach, the Jacobian matrix for the states, the controls and the constraints (Karush-Khun-Tucker Equations) while the boundaries are usually treated with the active set method (see Appendix C). Also for this method there is the problem of sub–optimal solutions.

Evolutionary Methods: also in this case continous problems needs to be discretized, but while direct methods could support a big number of variables the evolutionary ones cannot, usually they can deal with a small number of

unknowns for work effectively. The search is performed in many different ways, but all of them have one thing in common: they try to imitate the nature behaviour. The principal ones are:

- Genetic Algorithms: natural evolution is emulated, a set of individuals (the set of possible solutions) are randomly initialized, then their performance index is evaluated and a selection is performed taking this quantity as discriminant between who is suitable for the given ambient (the optimization problem) and who is not; after this selection the best individuals are used to generate a new generation of solutions and the algorithm is iterated until the value of a certain parameter (number of iterations, increment of performance index, etc.) is matched (this stop condition must be empirically fixed);

- Particle Swarm Optimization: this algorithm simulate fishes or birds looking for food, a randomly initialized set of solution is created and the individuals move in the domain of the problem. Their motion is assumed to be similar to the one adopted by a flock of animals when looking for food;

- Differential Evolution: also this method simulate the evolution, a randomly generated set of solutions is created but in this case the new population is created adding to the best individuals a component proportional to the difference between other two (or more than two) individuals.

Since the initialization is done randomly and the search method is not based on standard optimality conditions these methods should provide a better exploitation of the search space, but there is no proof of convergence.

1.4 Application of Perturbative Expansion to Low–Thrust Trajectory Arcs

From what it has been said before one could recognize two different trends:

1. low thrust (high I_{sp}) *adoption* as principal thrust supplier in interplanetary missions in order to save fuel;

2. deeper, more detailed and more effective optimization which can allow to reach in less time a more accurate globally optimal solution.

These are two requests which it is hard to accomplish togheter, in fact the introduction of low thrust in the equations of motion for a spacecraft make the system more complex fromt the orbit propagation point of view, so a higher computational cost must be added to the optimization process which become slower and less accurate respect to the case of chemical propulsion systems.

With this prospective the objective of this work it has been to look for some sufficiently accurate analytical approximations for the description of low thrust trajectory arcs and to demonstrate how to use this result in potentially more efficient optimization methods.

In particular for obtaining the analytical approximations two different methods has been developed:

- Method 1: starting from an application of the multiple scale perturbative approach found in the literature dealing with the motion of a spacecraft under the influence of a small drag force [11] and adapting the two dimensional equations of motion to the case of small accelaration, a first analytical result is obtained; the relations are able to describe the bidimensional motion and to capture only the secular variation for the orbital elements;

- Method 2: in order to obtain a simple three dimensional description of motion under low–thrust, Gauss's variational equations have been solved with a perturbative method, introducing the adimensionalized low thrust acceleration as perturbation parameter $\epsilon \ll 1$ and looking for a first order accurate solution.

A similar search for analytical solutions of low–thrust trajectory arcs has been already proposed in the literature. In some cases a particular trajectory shape is assumed, evaluating the correspondent thrust profile which exactly procuduce that trajectory [25], [15]. In other sudies the trajectory is derived without imposing a particular shape, trying to make the method as indipendent as possible from the particular thrust profile considered [18], [20].

In order to have an idea of the precision of the approximation derived with the perturbative approach presented in this thesis, a comparison with a method based on a Fourier series expansion [18] is presented in the results section. This method first mediates on one period Gauss's equations and obtains a mean rate of change for each orbital parameter, then evaluates the Fourier expansion for the given control profile and finally substitute it in the mediated equations; the analytical expression derived depends only on a small number of Fourier expansion coefficients.

As to the author's knowledge, a perturbation method was applied only in one case [7], for the approximation of the Hamiltonian H which derives from the classical formulation of the indirect optimization approach. In this respect, the novelty of the approach here developed lies in

- the possibility of propagating orbits under low–thrust propulsion with a very limited computational burden;
- the ability of the method in capturing the most relevant physical features of the perturbed solution;
- the derivation of simple criteria for evaluating the maximum arc length that results into an acceptable error;
- the availability of a simple low–thrust orbit discretization process suitable for applications in the framework of Direct Optimization Methods.

The most important limits of this method are

- the applicability of the expansions is limited to the space where the ratio between the thrust and the gravitational acceleration is smaller than 1;
- the error control requires to limit the maximum trajectory arc length, depending on the value of ε which in some case of practical interest could become equal to 10^{-1};

A preliminary application in the frame of Direct Optimization Methods has been outlined,in order to demonstrate the last point but it must be underlined that also for the evolutionary ones this analytical approximation could provide a significant improvement, spreading the set of possible problems which can be treated with them: if a big arc of low thrust trajectory can be solved analitically, the variables number for describing a low thrust transfer could become suitable for application of evolutionary algorithms.

As said before, Direct optimization methods transform the optimization problem into a multiple shooting problem: first of all, once initial and final states are given, intermediate values for the states and the controls must be choosen, in the very first run this is done without any kind of mathematical indication but only by means of the user ability and experience. The number of intermediate values depends on the accuracy required but it must be noticed that the number of variables is directly proportional to that quantity, so it cannot be too big in order to make the computation feasible. Besides the contrains on initial and final points and for values independent of the

trajectory and the discretization (e.g. limit on maximum thrust), another set of contrains must be enforced and these are the matching conditions at the interfaces for the sub arcs between the intermediate points, since the values are chosen indipendently for the various points and, in general the final point of one arc do not coincide with the initial one on the following.

The analytical approximations can be introduced for evaluating the trajectory in each sub arc instead of more computationally demanding numerical integration scheme and this is the most direct application. Since also the time of flight is evaluated analytically, also the Jacobian Matrix could be evaluated analytically by means of Automatic Differentiation, with a significant reduction of computational time. Obviously the results of the optimization process will be affected by the approximation inherent in the perturbation epxpansion but this approach could be used in order to identify a good first guess for a fully numerical optimization method.

The boudaries/limits of the reserch has been chosen to be

- restricted two body problem, all the perturbations (drag, solar radiation pressure, planet asphericity, etc.) are neglected;
- constant thrust magnitude and direction for deriving the analytical approximations for the motion.

  A radial-transverse-normal reference frame for the thrust components will be assumed;

The results would be used to find a good initial guess for a direct method, so that the restricted two body assumption could be considered good in this preliminary phase and it justify the consideration of only gravitational force; in addition the direct method automatically discretized the trajectory and the thrust magnitude and direction, i.e. control variables, are assumed constant in each, small, arch; so that these limitations are not so influent.

1.5 Thesis Structure

After this introduction, in Chapter Two is presented the core of the whole work: the theorical instruments are applied in real systems and a set of analytical approximations for describing the motion of a spacecraft under the influence of a small perturbation are derived; a possible application of these equations to Direct Optimization Methods by the use of Direct Transcription is reported in the last section. In Chapter Three the results are presented and in Chapter Four some conclusive remarks are reported. In Appendix A some fundamentals aspects of Space Flight Dynamics are recalled, in Appendix B

there are some basic concepts of Perturbation Methods and in Appendix C a general overview of Optimization Methods is outlined.

Chapter 2

Perturbation Methods on Orbital Parameters Variation

In this chapter the search for analytical relations for low–thrust trajectory arcs by means of the perturbative approach is performed. The two methods briefly outlined in the introduction are here applied to orbital problems: a two scale perturbation approach is developed first and the method based on standard perturbation follows. The results are two different set of relations:

- The first one, obtained by means of the two scale method, is able to capture the secular variation (long period terms) of three orbital parameters (a, e and ω), since the solution for the perturbative term is available only for the planar case;

- The second set of equations, obtained by means of a standard perturbation approach, is able to capture both the secular variation and the short period one for all the six orbit parameters, so it is a fully three dimensional formulation, and since it has been derived for the equinotial elements it is not affected by any kind of singularity.

Both results provide a new insight in the problem of the description of the behaviour of a spacecraft under the action of a low–thrust propulsion system. The second approach, in particular, represents a fully general method that can lend itself to be used as a novel orbit propagation method, also in the framework of low–thrust orbit transfer optimization problems, as it will be shown at the end of the next chapter.

2.1 Two Scale Perturbation

The classical example used for describing the two scale perturbation method is, as discussed in more details in Appendix B, the weakly nonlinear oscillator. It has been shown that is characterized by two different time scales and a particular approach is needed in order to capture the whole motion composed by the sum of the motions in these two different scales.

In nature there are a lot of examples of weakly nonlinear oscillators, in particular, as far as space flight mechanics problems are concerned, an interesting one is presented in [11]: the motion of a satellite moving under the influence of a primary body (i.e. a simple two body problem) affected by an elementary model of atmospheric drag ($\rho = cost$ and $D = f(V^2)$, $\vec{\boldsymbol{D}} \parallel \vec{\boldsymbol{V}}$). It can be shown that the resulting dynamics can be rappresented as a weakly nonlinear oscillator, if a 2D description of the motion is adopted, the orbit plane not being affected by drag which is an in–plane force.

Three dimensional examples are described in [12] and in [13], where a satellite moving around an oblate planet and in a restricted three–body problem is considered respectively. In both cases, the motion can be rappresented as a multi degree–of–freedom weakly nonlinear oscillator.

In what follows only the two–dimensional case will be addressed.

2.1.1 Thrust along Velocity Direction

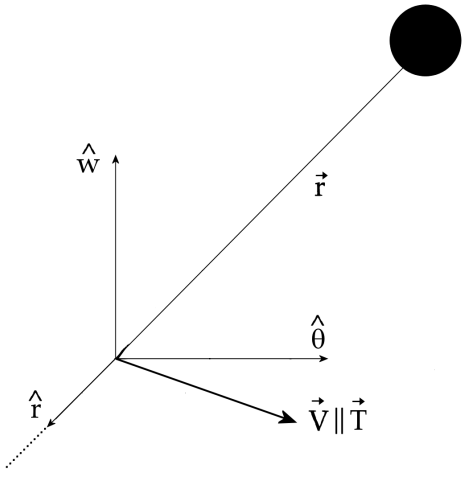

Figure 2.1: Thrust orientation

For the planar motion the equations are similar to those which describe a satellite affected by drag reported in [11]. The relations differ for the sign

changed from minus to plus in order to take into account the thrust instead of the drag; this change does not allow for directly finding the equations for the orbital parameters because of the nonlinearity of the problem. Thus, the equations are

$$\begin{aligned} m\left[\frac{d^2R}{dT^2} - R\left(\frac{d\theta}{dT}\right)^2\right] &= -\frac{GMm}{R^2} + F_0 f_1 \\ m\left(R\frac{d^2\theta}{dT^2} + 2\frac{dR}{dT}\frac{d\theta}{dT}\right) &= F_0 f_2 \end{aligned}$$

Scaling all the quantities with respect to a reference length $L_{ref} = R_0$ and a reference time $T_{ref} = (L^3/GM)^{1/2}$ one obtains the following pair of coupled, second order ordinary differential equations

$$\begin{aligned} \ddot{r} - r\dot{\theta}^2 &= -\frac{1}{r^2} + \varepsilon f_1\left(r, \theta, \dot{r}, \dot{\theta}\right) \\ r\ddot{\theta} + 2\dot{r}\dot{\theta} &= \varepsilon f_2\left(r, \theta, \dot{r}, \dot{\theta}\right) \end{aligned}$$

where $\varepsilon = F_0/(GmM/L^2) \ll 1$. Assuming that thrust acts along the velocity direction, the thrust vector can be decomposed in the radial transverse directions as follows:

$$F_0 f_1 = T\sin\gamma; \quad F_0 f_2 = T\cos\gamma$$

with

$$\tan\gamma = \frac{dR/dT}{Rd\theta/dT}$$

Since the velocity can be computed as

$$V = \left[\left(\frac{dR}{dT}\right)^2 + R^2\left(\frac{d\theta}{dT}\right)^2\right]^{1/2}$$

the equations of motion can be rewritten in dimensionless variables as

$$\begin{aligned} \ddot{r} - r\dot{\theta}^2 &= -\frac{1}{r^2} + \varepsilon\frac{\dot{r}}{\left(\dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2}} \\ r\ddot{\theta} + 2\dot{r}\dot{\theta} &= \varepsilon\frac{r\dot{\theta}}{\left(\dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2}} \end{aligned}$$

A more convenient form of the equations is obtained substituting the new variables $u(\theta,\varepsilon)$ and $t(\theta,\varepsilon)$ where $u \equiv 1/r$ and letting $' = d/d\theta$. With application chain rule,

$$\frac{d}{dt}(\bullet) = \frac{d}{d\theta}(\bullet)\frac{d\theta}{dt} = \frac{d}{d\theta}(\bullet)\frac{1}{t'}$$

one obtains

$$\begin{aligned}
\dot{r} &= -\frac{u'}{t'u^2} \\
\ddot{r} &= \frac{u'\left(u^2t'\right)' - u''u^2t'}{t'\left(t'u^2\right)^2} \\
\dot{\theta} &= \frac{1}{t'} \\
\ddot{\theta} &= -\frac{t''}{t'^3}
\end{aligned}$$

and the equations of motion become

$$\begin{aligned}
u'' + u - u^4t'^2 &= -\varepsilon u^2t'^2\left(f_1 + \frac{u'}{u}f_2\right) \\
\left(u^2t'\right)' &= -\varepsilon u^3t'^3f_2
\end{aligned}$$

Assuming that thrust is aligned with the velocity direction, the functions f_1 and f_2 are given by

$$\begin{aligned}
f_1 &= \frac{\dot{r}}{\left(\dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2}} \\
f_2 &= \frac{r\dot{\theta}}{\left(\dot{r}^2 + r^2\dot{\theta}^2\right)^{1/2}}
\end{aligned}$$

When expressed in terms of the new variables

$$\begin{aligned}
f_1 &= -\frac{u'}{(u'^2 + u^2)^{1/2}} \\
f_2 &= \frac{u}{(u'^2 + u^2)^{1/2}}
\end{aligned}$$

one gets the following formulation for the equations of motion:

$$\begin{aligned}
u'' + u - u^4t'^2 &= 0 \\
\left(u^2t'\right)' &= -\varepsilon\frac{u^4t'^3}{(u'^2 + u^2)^{1/2}}
\end{aligned} \tag{2.1}$$

Without loss of generality the initial conditions can be chosen as

$$\begin{aligned} u(0;\varepsilon) &= 1; \\ t(0;\varepsilon) &= 0; \\ u'(0;\varepsilon) &= 0; \\ u'(0;\varepsilon) &= \sigma; \end{aligned}$$

where σ is the reciprocal of the initial angular velocity, which must lie in the interval $2^{-1/2} < \sigma < 1$ [11][1]. These equations do not depend directly on t, and a two scale approximation can be introduced for the two functions u and t' following the preocedure proposed in [11][2] for the atmospheric drag case:

$$\begin{aligned} u(\theta;\varepsilon) &= u_0(\theta;\vartheta) + \varepsilon u_1(\theta;\vartheta) + ... \\ t'(\theta;\varepsilon) &= v_0(\theta;\vartheta) + \varepsilon v_1(\theta;\vartheta) + ... \end{aligned}$$

where $\vartheta = \varepsilon\theta$. The system thus becomes

$$\begin{aligned} u'' + u - u^4 t'^2 &= 0 \\ \left(u'^2 + u^2\right)\left(u^2 t'\right)' &= -\varepsilon\left(u^4 t'^3\right)\left(u'^2 + u^2\right)^{1/2} \end{aligned}$$

For the zero order terms it is

$$\begin{aligned} \frac{\partial^2 u_0}{\partial\theta^2} + u_0 &= u_0^4 v_0^2 \\ u_0^2 \frac{\partial v_0}{\partial\theta} + 2u_0 \frac{\partial u_0}{\partial\theta} v_0 &= 0 \end{aligned}$$

and the system solution is given by

$$\begin{aligned} u_0(\theta;\vartheta) &= \frac{1}{r} = p^2\left(1 + e\cos(\theta - \omega)\right) \\ v_0(\theta;\vartheta) &= \frac{1}{\vartheta} = p^{-3}\left(1 + e\cos(\theta - \omega)\right)^{-2} \end{aligned}$$

obviously identical to the one reported in [11].

[1] p. 329.
[2] p. 330.

For the first–order terms it is necessary to subsitute the first derivatives of u and v which can be expressed as

$$\begin{aligned} u' &= \frac{\partial u_0}{\partial \theta} + \varepsilon \left(\frac{\partial u_1}{\partial \theta} + \frac{\partial u_0}{\partial \vartheta} \right) \\ v' &= \frac{\partial v_0}{\partial \theta} + \varepsilon \left(\frac{\partial v_1}{\partial \theta} + \frac{\partial v_0}{\partial \vartheta} \right) \end{aligned}$$

After some algebraic manipulations, the following equations are obtained:

$$\begin{aligned} \frac{\partial^2 u_1}{\partial \theta^2} + u_1 &= -2\frac{\partial^2 u_0}{\partial \theta \partial \vartheta} + 2u_0^4 v_0 v_1 + 4u_0^3 u_1 v_0^2 \\ u_0^2 \left(\frac{\partial v_1}{\partial \theta} + \frac{\partial v_0}{\partial \vartheta} \right) + 2u_1 v_0 \frac{\partial u_0}{\partial \theta} &+ \\ +2u_0 \left[u_1 \frac{\partial v_0}{\partial \theta} + \frac{\partial u_0}{\partial \theta} v_1 + v_0 \left(\frac{\partial u_1}{\partial \theta} + \frac{\partial u_0}{\partial \bar{\theta}} \right) \right] &= -\frac{u_0^4 v_0^3}{\left(\left(\frac{\partial u_0}{\partial \theta} \right)^2 + u_0^2 \right)^{1/2}} \end{aligned} \tag{2.2}$$

The second equation can be written as

$$\frac{\partial}{\partial \theta} \left(u_0^2 v_1 + 2p \frac{u_1}{u_0} \right) + \frac{\partial}{\partial \vartheta} \left(v_0 u_0^2 \right) = -\frac{u_0^4 v_0^3}{\sqrt{u_0^2 + \frac{\partial u_0}{\partial \theta}^2}} \tag{2.3}$$

Introducing the partial derivatives

$$\begin{aligned} \frac{\partial u_0}{\partial \theta} &= p^2 \left[-e \sin \left(\theta - \omega \right) \right] \\ \frac{\partial v_0}{\partial \theta} &= p^{-3} \left[2 \left[1 + e \cos \left(\theta - \omega \right) \right]^{-3} e \sin \left(\theta - \omega \right) \right] \end{aligned}$$

and modifing the second term on the left side as

$$\frac{\partial}{\partial \theta} \left(u_0^2 v_1 + 2p \frac{u_1}{u_0} \right) + p' = -\frac{1}{p^3 \left[1 + e \cos \left(\theta - \omega \right) \right]^2 \left[1 + e^2 + 2e \cos(\theta - \omega) \right]^{1/2}}$$

a Taylor series expansion of the right side term can be considered

$$-\left\{ \left[1 + e \cos \left(\theta - \omega \right) \right]^2 \left[1 + e^2 + 2e \cos(\theta - \omega) \right]^{1/2} \right\}^{-1} \cong 3e \cos \left(\theta - \omega \right) - 1 + \mathcal{O}\left(e^2 \right)$$

so that Eq. (2.3) can be rewritten in the form

$$\frac{\partial}{\partial \theta} \left(u_0^2 v_1 + 2p \frac{u_1}{u_0} \right) + p' = \frac{1}{p^3} \left(3e \cos \left(\theta - \omega \right) - 1 \right)$$

which can be integrated once with respect to θ:

$$u_0^2 v_1 + 2p\frac{u_1}{u_0} + p'\theta - \frac{1}{p^3}\left(3e\sin\left(\theta-\omega\right)-\theta\right) = g_1\left(\vartheta\right) + \mathcal{O}\left(e^2\right)$$

The first equation in (2.2) can be rewritten as

$$\frac{\partial^2 u_1}{\partial\theta^2} + u_1 = -2\frac{\partial^2 u_0}{\partial\theta\partial\vartheta} + 2u_0^2 v_0\left(u_0^2 v_1 + 2p\frac{u_1}{u_0}\right) \tag{2.4}$$

and the last term can be found from the equation (2.4), so, since

$$\frac{\partial}{\partial\vartheta}\left(\frac{\partial u_0}{\partial\theta}\right) = -\left(e'p^2 + 2epp'\right)\sin\left(\theta-\omega\right) + \omega' ep^2\cos\left(\theta-\omega\right)$$

and $2u_0^2 v_0 = 2p$, Eq (2.4) becomes

$$\begin{aligned}\frac{\partial^2 u_1}{\partial\theta^2} + u_1 &= \left(2p^2\frac{\partial e}{\partial\vartheta} + 4ep\frac{\partial p}{\partial\vartheta} + \frac{6e}{p^2}\right)\sin\left(\theta-\omega\right) - 2ep^2\frac{\partial\omega}{\partial\vartheta}\cos\left(\theta-\omega\right) + \\ &- 2\left(p\frac{\partial p}{\partial\vartheta} + \frac{1}{p^2}\right)\theta + 2pg_1\end{aligned}$$

The secular term vanishes provided that the following conditions are satisfied:

$$\frac{\partial\omega}{\partial\vartheta} = 0 \Rightarrow \omega = cost = \omega_0$$

$$p^3\frac{\partial p}{\partial\vartheta} = -1 \Rightarrow a = a_0\sqrt{\frac{1}{1-4a_0^2\varepsilon\theta}}$$

$$2p^2\frac{\partial e}{\partial\vartheta} - \frac{4ep}{p^3} + \frac{6e}{p^2} = 2p^2\frac{\partial e}{\partial\vartheta} + \frac{2e}{p^2} = 0$$

that is

$$\frac{\partial e}{e} = \frac{\partial\vartheta}{4\vartheta - p_0^4} \Rightarrow e = e_0\sqrt[4]{1-4a_0^2\varepsilon\theta}$$

It must be underlined that the final relations derived for the orbital parameters evolution are different from the relations derived in the drag case, and this difference is not a simple sign change. This is a direct consequence of the nonlinearity of the system.

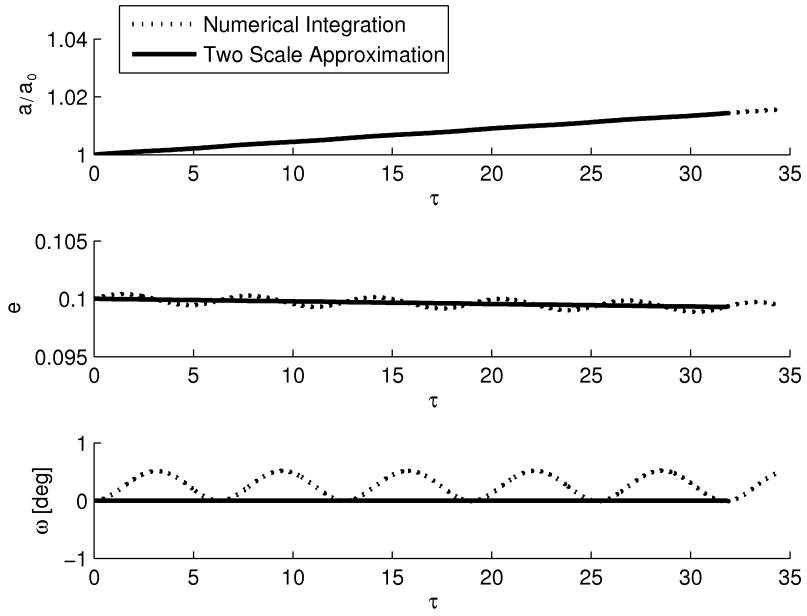

Figure 2.2: Thrust along Velocity, ($\varepsilon = 10^{-3}$)

2.1.2 Constant low–thrust case

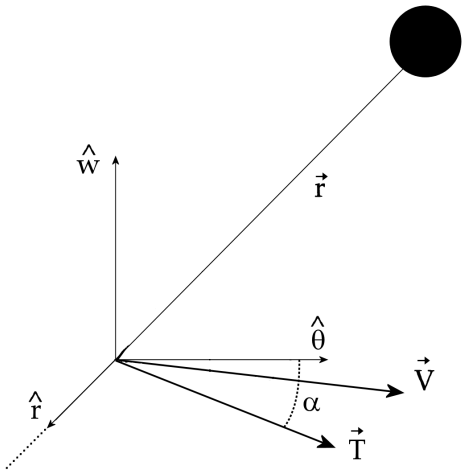

Figure 2.3: Thrust orientation

If a thrust vector along a fixed direction with respect to the radial-transversal reference system is assumed, where α is the angle between thrust and transversal direction, it can be decomposed as

$$F_0 f_1 = T \sin\alpha; \quad F_0 f_2 = T \cos\alpha$$

The equations of motion in dimensionless form are given by

$$\begin{aligned} \ddot{r} - r\dot{\theta}^2 &= -\frac{1}{r^2} + \varepsilon \sin\alpha \\ r\ddot{\theta} + 2\dot{r}\dot{\theta} &= \varepsilon \cos\alpha \end{aligned}$$

Following similar steps to those presented for the athmospheric drag case the following relation was obtained

$$\begin{aligned} \frac{\partial^2 u_1}{\partial\theta^2} + u_1 &= -2\left[-\left(e'p^2 + 2epp'\right)\sin(\theta-\omega) + \omega' e p^2 \cos(\theta-\omega)\right] + \\ &+ 2p\left[-p'\theta - \frac{\cos\alpha}{p^3}\left(\theta - 3e\sin(\theta-\omega)\right) + g_1(\vartheta)\right] + \\ &+ \left[\frac{e\sin(\theta-\omega)\cos\alpha}{p^2\left[1+e\cos(\theta-\omega)\right]^3} - \frac{\sin\alpha}{p^2\left[1+e\cos(\theta-\omega)\right]^2}\right] \end{aligned} \tag{2.5}$$

Introducing the following Taylor series

$$\begin{aligned}\frac{e\sin(\theta-\omega)}{[1+e\cos(\theta-\omega)]^3} &\cong e\sin(\theta-\omega)+\mathcal{O}\left(e^2\right)\\ -\frac{1}{[1+e\cos(\theta-\omega)]^2} &\cong 2e\cos\left(\theta-\omega\right)-1+\mathcal{O}\left(e^2\right)\end{aligned}$$

Eq (2.5) becomes

$$\begin{aligned}\frac{\partial^2 u_1}{\partial\theta^2}+u_1 &= \left(2p^2\frac{\partial e}{\partial\vartheta}+4ep\frac{\partial p}{\partial\vartheta}+\frac{7e}{p^2}\cos\alpha\right)\sin\left(\theta-\omega\right)+\\ &+ \left(-2ep^2\frac{\partial\omega}{\partial\vartheta}+\frac{2e}{p^2}\sin\alpha\right)\cos\left(\theta-\omega\right)+\\ &+ 2p\left(-\frac{\partial p}{\partial\vartheta}-\frac{\cos\alpha}{p^3}\right)\theta-\frac{\sin\alpha}{p^2}+2pg_1\end{aligned}$$

Again, making the secular terms vanish one gets

$$p^3\frac{\partial p}{\partial\vartheta}=-\cos\alpha\Rightarrow a=a_0\sqrt{\frac{1}{1-4\cos\alpha a_0^2\varepsilon\theta}}$$

$$\frac{\partial\omega}{\partial\vartheta}=\frac{\sin\alpha}{p^4}\Rightarrow$$

$$\Rightarrow\begin{cases}\omega &= \frac{\tan\alpha}{4}\ln\left(\frac{1}{1-4\cos\alpha a_0^2\varepsilon\theta}\right) & \text{if}\quad \alpha\neq\frac{\pi}{2}\\ \omega &= a_0^2\varepsilon\theta & \text{if}\quad \alpha=\frac{\pi}{2}\end{cases}$$

$$2p^2\frac{\partial e}{\partial\vartheta}-\frac{4ep}{p^3}\cos\alpha+\frac{7e}{p^2}\cos\alpha=2p^2\frac{\partial e}{\partial\vartheta}+\frac{3e}{p^2}\cos\alpha=0$$

$$\frac{\partial e}{e}=\frac{3}{2}\cos\alpha\frac{\partial\vartheta}{4\cos\alpha\vartheta-p_0^4}\Rightarrow e=e_0\left(1-4\cos\alpha a_0^2\varepsilon\theta\right)^{\frac{3}{8}}$$

It must be underlined that these expansions work both forward and backward in angular position, simply adding the minus sign inside the value of ϑ.

In what follows are reported three figures (2.4, 2.5 and 2.6) for three different thrust orientation angles. The approximation is compared with an accurate numerical solution, as can be seen the results are qualitatively similar for the different thrust orientation. The analytical approximation is not able to capture the short term variation but the secular trend is catched quite accurately.

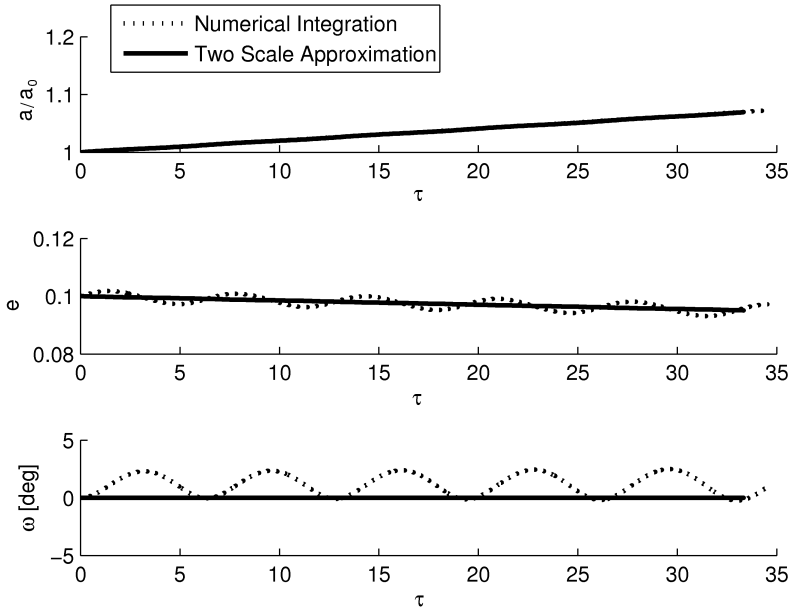

Figure 2.4: $\alpha = 0$ deg, $(\varepsilon = 10^{-3})$

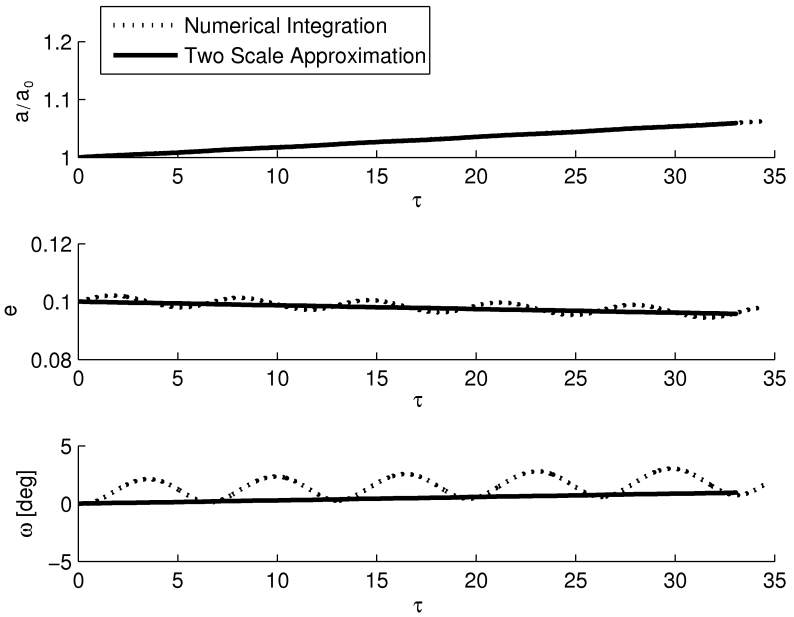

Figure 2.5: $\alpha = 30$ deg, $(\varepsilon = 10^{-3})$

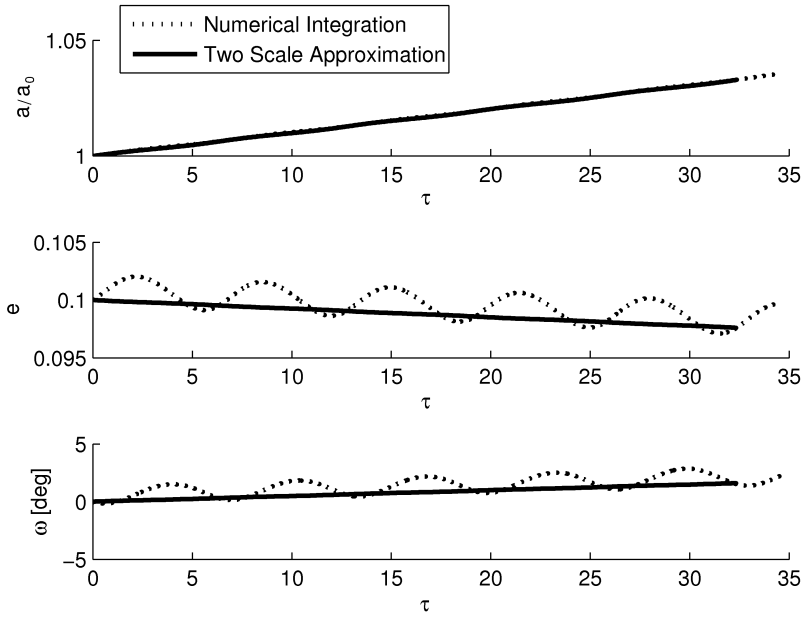

Figure 2.6: $\alpha = 60$ deg, $(\varepsilon = 10^{-3})$

2.2 Standard Perturbation

2.2.1 Classical Parameter Formulation

For these analysis a constant thrust $\vec{\boldsymbol{T}}$ is assumed, its orientation with respect to the radial direction $\hat{\boldsymbol{r}}$, the trasverse one $\hat{\boldsymbol{\theta}}$ and normal to the orbit plane $\hat{\boldsymbol{w}}$ is defined by means of the following angles: α between the radial direction and the projection of the thrust vector on the $\hat{\boldsymbol{r}} - \hat{\boldsymbol{\theta}}$ plane, and β between the thrust vector and the $\hat{\boldsymbol{r}} - \hat{\boldsymbol{\theta}}$ plane (cf. fig 2.7). With this decomposition

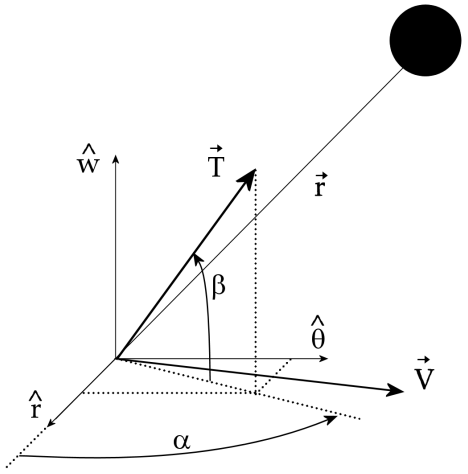

Figure 2.7: Thrust orientation

one can define the following quantities:

$$T_w = T \sin\beta; \quad T_r = T \cos\beta \cos\alpha; \quad T_\theta = T \cos\beta \sin\alpha$$

Thrust magnitude is made non–dimensional scaling it respect to the vechicle weight $\mu m/r^2$, that is a perturbation parameter

$$\varepsilon = \frac{T}{\mu m/r^2}$$

can be defined, which can be obviously seen as the ratio between the acceleration caused by thrust and that caused by the gravitational attraction . The semimajor axis a_0 of the initial orbit will be assumed as the reference length, while a mass parameter $\mu = 1$ is also assumed, such that the orbit period in non–dimensional terms is $\tau = 2\pi$ and the following derivation is indipendent of the particular celestial body considered.

Considering now Gauss's equations for variation of orbit parameters (reported in Appendix A)

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{r\sin(\nu+\omega)}{h\sin i}a_w \\
\frac{di}{dt} &= \frac{r\cos(\nu+\omega)}{h}a_w \\
\frac{d\omega}{dt} &= \frac{1}{he}\left[-p\cos\nu a_r + (p+r)\sin\nu a_\theta\right] - \frac{d\Omega}{dt}\cos i \\
\frac{da}{dt} &= \frac{2a^2}{h}\left(e\sin\nu a_r + \frac{p}{r}a_\theta\right) \\
\frac{de}{dt} &= \frac{1}{h}\left\{p\sin\nu a_r + \left[(p+r)\cos\nu + re\right]a_\theta\right\} \\
\frac{dM}{dt} &= n + \frac{b}{ahe}\left[(p\cos\nu - 2re)\,a_r - (p+r)\sin\nu a_\theta\right]
\end{aligned}$$

where the acceleration components a_i, $i = w, r, \theta$ can be replaced by the corresponding terms produced by the low–thrust propulsion system that is $a_i = T_i/m$.

Replacing derivation with respecto to time with derivation with respect to the true anomaly by application of the chain rule

$$\frac{d}{dt} = \frac{d}{d\nu}\frac{d\nu}{dt} \Rightarrow \frac{d}{d\nu} = \frac{d}{dt}\frac{1}{\dot{\nu}}$$

one can rewrite the system of first order ordinary differential equations as follows:

$$\begin{aligned}
\frac{d\Omega}{d\nu} &= \frac{r^3\sin(\nu+\omega)}{h^2\sin i}\varepsilon\sin\beta \\
\frac{di}{d\nu} &= \frac{r^3\cos(\nu+\omega)}{h^2}\varepsilon\sin\beta \\
\frac{d\omega}{d\nu} &= \frac{r^2}{h^2e}\left[-p\cos\nu\varepsilon\cos\beta\cos\alpha + (p+r)\sin\nu\varepsilon\cos\beta\sin\alpha\right] - \frac{d\Omega}{d\nu}\cos i \\
\frac{da}{d\nu} &= \frac{2r^2a^2}{h^2}\left(e\sin\nu\varepsilon\cos\beta\cos\alpha + \frac{p}{r}\varepsilon\cos\beta\sin\alpha\right) \\
\frac{de}{d\nu} &= \frac{r^2}{h^2}\left\{p\sin\nu\varepsilon\cos\beta\cos\alpha + \left[(p+r)\cos\nu + re\right]\varepsilon\cos\beta\sin\alpha\right\} \\
\frac{dM}{d\nu} &= \frac{r^2}{a\sqrt{ap}} + \frac{\sqrt{ap}r^2}{ah^2e}\left[(p\cos\nu - 2re)\,\varepsilon\cos\beta\cos\alpha - (p+r)\sin\nu\varepsilon\cos\beta\sin\alpha\right]
\end{aligned}$$

where the quantities

$$\begin{aligned} b &= \sqrt{pa} \\ h^2 &= p\mu \\ n &= h/ab \\ \dot{\nu} &= h/r^2 \end{aligned}$$

are used.

Then, introducing

$$\begin{aligned} r &= \frac{p}{[1 + e\cos\nu]} \\ p &= a\left(1 - e^2\right) \\ h^2 &= \frac{p}{\mu} \end{aligned}$$

in the last system, it becomes

$$\begin{aligned}
\frac{d\Omega}{d\nu} &= \frac{a^2\left(1-e^2\right)^2\sin\left(\nu+\omega\right)}{\left[1+e\cos\left(\nu\right)\right]^3\sin i}\varepsilon\sin\beta \\
\frac{di}{d\nu} &= \frac{a^2\left(1-e^2\right)^2\cos\left(\nu+\omega\right)}{\left[1+e\cos\left(\nu\right)\right]^3}\varepsilon\sin\beta \\
\frac{d\omega}{d\nu} &= \frac{a\left(1-e^2\right)}{\left[1+e\cos\nu\right]^2 e}\left[-a\left(1-e^2\right)\cos\nu\varepsilon\cos\beta\cos\alpha+\right. \\
&+ \left. a\left(1-e^2\right)\left(1+\frac{1}{\left[1+e\cos\nu\right]}\right)\sin\nu\varepsilon\cos\beta\sin\alpha\right]-\frac{d\Omega}{d\nu}\cos i \\
\frac{da}{d\nu} &= \frac{2a^3\left(1-e^2\right)}{\left[1+e\cos\nu\right]^2}\left(e\sin\nu\varepsilon\cos\beta\cos\alpha+\left[1+e\cos\nu\right]\varepsilon\cos\beta\sin\alpha\right) \\
\frac{de}{d\nu} &= \frac{a\left(1-e^2\right)}{\left[1+e\cos\nu\right]^2}\left\{a\left(1-e^2\right)\sin\nu\varepsilon\cos\beta\cos\alpha+\right. \\
&+ \left[a\left(1-e^2\right)\left(1+\frac{1}{\left[1+e\cos\nu\right]}\right)\cos\nu+\right. \\
&+ \left.\left.\frac{ae(1-e^2)}{\left[1+e\cos\nu\right]}\right]\varepsilon\cos\beta\sin\alpha\right\} \\
\frac{dM}{d\nu} &= \frac{\sqrt{1-e^2}}{\left[1+e\cos\nu\right]^2}+\frac{\sqrt{(1-e^2)}a\left(1-e^2\right)}{e\left[1+e\cos\nu\right]^2}\left[a\left(1-e^2\right)\times\right. \\
&\times \left(\cos\nu-\frac{2e}{\left[1+e\cos\nu\right]}\right)\varepsilon\cos\beta\cos\alpha+ \\
&- \left. a\left(1-e^2\right)\left(1+\frac{1}{\left[1+e\cos\nu\right]}\right)\sin\nu\varepsilon\cos\beta\sin\alpha\right]
\end{aligned}$$

At this point the next step is to introduce the perturbative expansions for all the orbital parameters and to substitute these relations into the differential equation system. The expansions are truncated to first–order terms, so they are written as

$$\begin{aligned}
a &= a_0+\varepsilon a_1 \\
e &= e_0+\varepsilon e_1 \\
i &= i_0+\varepsilon i_1 \\
\omega &= \omega_0+\varepsilon\omega_1 \\
\Omega &= \Omega_0+\varepsilon\Omega_1 \\
M &= M_0+\varepsilon M_1
\end{aligned}$$

Introducing them into the previous differential equation system (the symbol ' denote the derivation with respect to ν) and expressing only those factors that give rise to first–order terms one obtains

$$\begin{aligned}
\Omega_0' + \varepsilon\Omega_1' &= \frac{a_0^2\left(1-e_0^2\right)^2 \sin\left(\nu+\omega_0\right)}{\left[1+e_0\cos\left(\nu\right)\right]^3 \sin i_0}\varepsilon\sin\beta \\
i_0' + \varepsilon i_1' &= \frac{a_0^2\left(1-e_0^2\right)^2 \cos\left(\nu+\omega_0\right)}{\left[1+e_0\cos\left(\nu\right)\right]^3}\varepsilon\sin\beta \\
\omega_0' + \varepsilon\omega_1' &= \frac{a_0\left(1-e_0^2\right)}{\left[1+e_0\cos\nu\right]^2 e_0}\left[-a_0\left(1-e_0^2\right)\cos\nu\varepsilon\cos\beta\cos\alpha+\right. \\
&+ \left. a_0\left(1-e_0^2\right)\left(1+\frac{1}{\left[1+e_0\cos\nu\right]}\right)\sin\nu\varepsilon\cos\beta\sin\alpha\right]+ \\
&- \Omega_0'\cos i_0 - \varepsilon\Omega_1'\cos i_0 \\
a_0' + \varepsilon a_1' &= \frac{2a_0^3\left(1-e_0^2\right)}{\left[1+e_0\cos\nu\right]^2}\left(e_0\sin\nu\varepsilon\cos\beta\cos\alpha+\left[1+e_0\cos\nu\right]\varepsilon\cos\beta\sin\alpha\right) \\
e_0' + \varepsilon e_1' &= \frac{a_0\left(1-e_0^2\right)}{\left[1+e_0\cos\nu\right]^2}\left\{a_0\left(1-e_0^2\right)\sin\nu\varepsilon\cos\beta\cos\alpha+\right. \\
&+ \left.\left[a_0\left(1-e_0^2\right)\left(1+\frac{1}{\left[1+e_0\cos\nu\right]}\right)\cos\nu+\frac{a_0e_0(1-e_0^2)}{\left[1+e_0\cos\nu\right]}\right]\varepsilon\cos\beta\sin\alpha\right\} \\
M_0' + \varepsilon M_1' &= \frac{\sqrt{1-e^2}}{\left[1+e\cos\nu\right]^2}+\frac{a_0\left(1-e_0^2\right)^{3/2}}{e_0\left[1+e_0\cos\nu\right]^2}\left[a_0\left(1-e_0^2\right)\left(\cos\nu-\frac{2e_0}{\left[1+e_0\cos\nu\right]}\right)\times\right. \\
&\times \left.\varepsilon\cos\beta\cos\alpha-a_0\left(1-e_0^2\right)\left(1+\frac{1}{\left[1+e_0\cos\nu\right]}\right)\sin\nu\varepsilon\cos\beta\sin\alpha\right]
\end{aligned}$$

One can note that, similarly to what will be shown in details for equinotial elements, once ν is known the information given by the last quantity M is redudant, i.e. not stricly necessary for defining vehicle position. This fact apparently allows one to ignore one equation with no loss of information, but this is not the case: another equation is needed to close the system in time, which si derived from that used to perform the change in derivation variable by means of the chain rule, that is:

$$\frac{1}{\dot{\nu}} = \frac{r^2}{h}$$

With this strategy, one change the focus from time as indipendent variable to ν, so it is for this reason that the equation for M is apparently not needed anymore, but, in doing this, time becomes a dipendent variable, and an equation is needed for evaluating it. Using the previous relation and the

previous defined quantities (expressed as functions of ν) one can write

$$\frac{dt}{d\nu} = \frac{r^2}{h} = \frac{p^2}{h\mu\left(1 + e\cos\nu\right)^2}$$

which is an equation similar to the others, where one can introduce an approximate perturbative expansion for the time:

$$t = t_0 + \varepsilon t_1$$

Upon substitution, one can obtain the perturbed variation of time, that will be derived later in the chapter.

Solving for the terms of zero order one obtain

$$\begin{aligned}
a_0 &= \bar{a}_0 \\
e_0 &= \bar{e}_0 \\
i_0 &= \bar{i}_0 \\
\omega_0 &= \bar{\omega}_0 \\
\Omega_0 &= \bar{\Omega}_0 \\
M_0 &= \sqrt{1 - \bar{e}_0^2} \int_{\nu_0}^{\nu} \frac{1}{\left[1 + e_0\cos\nu\right]^2}
\end{aligned}$$

where the values with an overbar are referred to quantities assigned at the

initial time t_0. For the first–order terms one gets

$$
\begin{aligned}
\Omega_1 &= \frac{a_0^2\left(1-e_0^2\right)^2\sin\beta}{\sin i_0}\left[\cos\omega_0\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^3}+\sin\omega_0\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}\right]\\
i_1 &= a_0^2\left(1-e_0^2\right)^2\sin\beta\left[\cos\omega_0\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}-\sin\omega_0\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^3}\right]\\
\omega_1 &= -\frac{a_0^2\left(1-e_0^2\right)^2\cos\beta\cos\alpha}{e_0}\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^2}+\\
&+ \frac{a_0^2\left(1-e_0^2\right)^2\cos\beta\sin\alpha}{e_0}\left[\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^2}+\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^3}\right]\\
&- \left(\Omega_1-\Omega_{1_0}\right)\cos i0\\
a_1 &= 2a_0^3\left(1-e_0^2\right)e_0\cos\beta\cos\alpha\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^2}+\\
&+ 2a_0^3\left(1-e_0^2\right)\cos\beta\sin\alpha\int_{\nu_0}^{\nu}\frac{1}{\left[1+e_0\cos\nu\right]}\\
e_1 &= a_0^2\left(1-e_0^2\right)^2\cos\beta\cos\alpha\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^2}+\\
&+ a_0^2\left(1-e_0^2\right)^2\cos\beta\sin\alpha\left[\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^2}+\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}\right]+\\
&+ a_0^2(1-e_0^2)^2e_0\cos\beta\sin\alpha\int_{\nu_0}^{\nu}\frac{1}{\left[1+e_0\cos\nu\right]^3}
\end{aligned}
$$

The last term is a bit trickier than the other ones as in this latter case one must take into account also possible first–order terms that derive from the zero order term, when substituting the perturbative expansions into it. In particular the zero order term becomes

$$
\frac{\sqrt{1-\left(e_0+\varepsilon e_1\right)^2}}{\left[1+\left(e_0+\varepsilon e_1\right)\cos\nu\right]^2}\Rightarrow
$$

$$
\Rightarrow\frac{e_1\left(e_0^2-2\right)}{\sqrt{1-e_0^2}}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}-\frac{e_1e_0}{\sqrt{1-e_0^2}}\frac{1}{\left[1+e_0\cos\nu\right]^3}=\mathcal{O}\left(\varepsilon\right)
$$

so that

$$\begin{aligned} M_1 &= \frac{a_0^2\left(1-e_0^2\right)^{5/2}}{e_0}\cos\beta\cos\alpha\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^2}+ \\ &- 2a_0^2\left(1-e_0^2\right)^{5/2}\cos\beta\cos\alpha\int_{\nu_0}^{\nu}\frac{1}{\left[1+e_0\cos\nu\right]^3}+ \\ &- \frac{a_0^2\left(1-e_0^2\right)^2}{e_0}\cos\beta\sin\alpha\left[\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^2}-\int_{\nu_0}^{\nu}\frac{\sin\nu}{\left[1+e_0\cos\nu\right]^3}\right] \\ &+ \frac{\left(e_0^2e_1-2e_1\right)}{\sqrt{1-e_0^2}}\int_{\nu_0}^{\nu}\frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}-\frac{e_0e_1}{\sqrt{1-e_0^2}}\int_{\nu_0}^{\nu}\frac{1}{\left[1+e_0\cos\nu\right]^3} \end{aligned}$$

Noting that

$$\begin{aligned} \frac{\sin\left(\nu+\omega_0\right)}{\left[1+e_0\cos\nu\right]^3} &= \frac{\sin\nu\cos\omega_0+\cos\nu\sin\omega_0}{\left[1+e_0\cos\nu\right]^3} \\ \frac{\cos\left(\nu+\omega_0\right)}{\left[1+e_0\cos\nu\right]^3} &= \frac{\cos\nu\sin\omega_0-\sin\nu\cos\omega_0}{\left[1+e_0\cos\nu\right]^3} \end{aligned}$$

the complete solution of the problem requires the solution of following inte-

grals

$$
\begin{aligned}
\int \frac{\sin\nu}{\left[1+e_0\cos\nu\right]^2}d\nu &= \frac{1}{e_0\left(1+e_0\cos\nu\right)} \\
\int \frac{\sin\nu}{\left[1+e_0\cos\nu\right]^3}d\nu &= \frac{1}{2e_0\left(1+e_0\cos\nu\right)^2} \\
\int \frac{1}{\left[1+e_0\cos\nu\right]}d\nu &= \frac{\nu}{\sqrt{1-e_0^2}} - \frac{2\arctan\left(\frac{2e_0\sin\nu}{2e_0\cos\nu+\left(\sqrt{1-e_0}+\sqrt{1+e_0}\right)^2}\right)}{\sqrt{1-e_0^2}} \\
\int \frac{1}{\left[1+e_0\cos\nu\right]^2}d\nu &= -\frac{2\arctan\left(\frac{2e_0\sin\nu}{2e_0\cos\nu+\left(\sqrt{1-e_0}+\sqrt{1+e_0}\right)^2}\right)}{\left(1-e_0^2\right)^{3/2}} + \\
&+ \frac{e_0\sin\nu}{\left(e_0^2-1\right)\left(1+e_0\cos\nu\right)} + \frac{\nu}{\left(1-e_0^2\right)^{3/2}} \\
\int \frac{1}{\left[1+e_0\cos\nu\right]^3}d\nu &= \left(\frac{1}{\left(1-e_0^2\right)^{3/2}} - \frac{3}{\left(1-e_0^2\right)^{5/2}}\right)\times \\
&\times \arctan\left(\frac{2e_0\sin\nu}{2e_0\cos\nu+\left(\sqrt{1-e_0}+\sqrt{e_0+1}\right)^2}\right) + \\
&- \frac{e_0\sin\nu\left(3e_0\cos\nu-e_0^2+4\right)}{2\left(e_0^2-1\right)^2\left(1+e_0\cos\nu\right)^2} - \frac{\nu}{2\left(1-e_0^2\right)^{3/2}} + \frac{3\nu}{2\left(1-e_0^2\right)^{5/2}} \\
\int \frac{\cos\nu}{\left[1+e_0\cos\nu\right]^2}d\nu &= \left(\frac{2}{e_0\left(1-e_0^2\right)^{3/2}} - \frac{2}{e_0\sqrt{1-e_0^2}}\right)\times \\
&\times \arctan\left(\frac{2e_0\sin\nu}{2e_0\cos\nu+\left(\sqrt{1-e_0}+\sqrt{e_0+1}\right)^2}\right) + \\
&+ \frac{\sin\nu}{\left(1-e_0^2\right)\left(1+e_0\cos\nu\right)} + \frac{\nu}{e_0\sqrt{1-e_0^2}} - \frac{\nu}{e_0\left(1-e_0^2\right)^{3/2}} \\
\int \frac{\cos\nu}{\left[1+e_0\cos\nu\right]^3}d\nu &= \left(\frac{e_0^2+2}{e_0\left(1-e_0^2\right)^{5/2}} - \frac{2}{e_0\left(1-e_0^2\right)^{3/2}}\right)\times \\
&\times \arctan\left(\frac{2e_0\sin\nu}{2e_0\cos\nu+\left(\sqrt{1-e_0}+\sqrt{e_0+1}\right)^2}\right) + \\
&+ \frac{\sin\nu\left(e_0\left(2e_0^2+1\right)\cos\nu+e_0^2+2\right)}{2\left(e_0^2-1\right)^2\left(1+e_0\cos\nu\right)^2} + \\
&+ \frac{3\nu}{2e_0\left(1-e_0^2\right)^{3/2}} - \frac{3\nu}{2e_0\left(1-e_0^2\right)^{5/2}}
\end{aligned}
$$

If one does not want to solve completely the integrals, which is comptutationally expensive, a Taylor series expansion with respect to eccentricity e

could be adopted, if low eccentricity orbits are dealt with. A first–order series results in an acceptable approximation if e is sufficiently small. A set of possible expansions is reported in the last section of this paragraph.

The Taylor series for the previous integrals are proposed in what follows

$$\begin{aligned}
\frac{\sin(\nu+\omega_0)}{[1+e_0\cos\nu]^3} &\cong \sin(\nu+\omega_0)(1-3e_0\cos\nu)+\mathcal{O}\left(e_0^2\right)\\
\frac{\cos(\nu+\omega_0)}{[1+e_0\cos\nu]^3} &\cong \cos(\nu+\omega_0)(1-3e_0\cos\nu)+\mathcal{O}\left(e_0^2\right)\\
\frac{\cos\nu}{[1+e_0\cos\nu]^2} &\cong \cos\nu-2e_0\cos^2\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{\sin\nu}{[1+e_0\cos\nu]^2} &\cong \sin\nu-2e_0\sin\nu\cos\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{\sin\nu}{[1+e_0\cos\nu]^3} &\cong \sin\nu-3e_0\sin\nu\cos\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{1}{[1+e_0\cos\nu]} &\cong 1-e_0\cos\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{\cos\nu}{[1+e_0\cos\nu]^3} &\cong \cos\nu-3e_0\cos^2\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{1}{[1+e_0\cos\nu]^3} &\cong 1-3e_0\cos\nu+\mathcal{O}\left(e_0^2\right)\\
\frac{\sin\nu}{[1+e_0\cos\nu]} &\cong \sin\nu-e_0\sin\nu\cos\nu+\mathcal{O}\left(e_0^2\right)
\end{aligned}$$

In order to obtain a formulation valid for thrust components projected along the velocity-normal frame, one can transform Gauss's equations by means of a rotation matrix that transforms the vectors expressed with respect to $\hat{\boldsymbol{r}}-\hat{\boldsymbol{\theta}}$ into frame components expressed with respect to the direction of the velocity and in–plane normal direction unit vectors:

$$\begin{bmatrix} a_{dr} \\ a_{d\theta} \end{bmatrix} = \frac{h}{pv} \begin{bmatrix} e\sin\nu & -(1+e\cos\nu) \\ 1+e\cos\nu & e\sin\nu \end{bmatrix} \begin{bmatrix} a_{dt} \\ a_{dn} \end{bmatrix}$$

thus obtaining the equations reported in [2]. Expressing v as a funtion of the orbit parameters one can apply the same perturbative technique.

2.2.2 Equinocial Orbital Parameters Formulation

In order to obtain a nonsingular formulation, the perturbation approach will now be applied to the modified equinocial orbit parameters. Given the fol-

lowing quantities

$$P_1 = e\sin\varpi \quad Q_1 = \tan\frac{i}{2}\sin\Omega$$
$$P_2 = e\cos\varpi \quad Q_2 = \tan\frac{i}{2}\cos\Omega$$

with

$$\varpi = \Omega + \omega \quad l = \varpi + M \quad L = \varpi + \nu;$$

the equations for the variation of equinocial parameters are

$$\begin{aligned}
\frac{da}{dt} &= \frac{2a^2}{h}\left[\left(P_2\sin L - P_1\cos L\right)a_r + \frac{p}{r}a_\theta\right] \\
\frac{dP_1}{dt} &= \frac{r}{h}\left\{-\frac{p}{r}\cos L a_r + \left[P_1 + \left(1+\frac{p}{r}\right)\sin L\right]a_\theta - P_2\left(Q_1\cos L - Q_2\sin L\right)a_w\right\} \\
\frac{dP_2}{dt} &= \frac{r}{h}\left\{\frac{p}{r}\sin L a_r + \left[P_2 + \left(1+\frac{p}{r}\right)\cos L\right]a_\theta + P_1\left(Q_1\cos L - Q_2\sin L\right)a_w\right\} \\
\frac{dQ_1}{dt} &= \frac{r}{2h}\left(1+Q_1^2+Q_2^2\right)\sin L a_w \\
\frac{dQ_2}{dt} &= \frac{r}{2h}\left(1+Q_1^2+Q_2^2\right)\cos L a_w \\
\frac{dl}{dt} &= n - \frac{r}{h}\left\{\left[\frac{a}{a+b}\left(\frac{p}{r}\right)\left(P_1\sin L + P_2\cos L\right) + \frac{2b}{a}\right]a_r\right. \\
&+ \left.\frac{a}{a+b}\left(1+\frac{p}{r}\right)\left(P_1\cos L - P_2\sin L\right)a_\theta + \left(Q_1\cos L - Q_2\sin L\right)a_w\right\}
\end{aligned}$$

Here, again, it is possible to substitute the acceleration terms with thrust components per unit mass and to express the derivative with respect to the true longitude instead of time, according to the following relation

$$\frac{d}{dt} = \frac{d}{dL}\frac{dL}{dt} \Rightarrow \frac{d}{dL} = \frac{d}{dt}\frac{1}{\dot{L}}$$

It should be noted that

$$\dot{L} = \dot{\Omega} + \dot{\omega} + \dot{\nu}$$

so that, when the perturbative expansions truncated at the first–order is used, one gets

$$\dot{L} = \dot{\Omega}_0 + \varepsilon\dot{\Omega}_1 + \dot{\omega}_0 + \varepsilon\dot{\omega}_1 + \dot{\nu}$$

Remembering that

$$\dot{\Omega}_0 = 0 \quad \dot{\omega}_0 = 0$$

and neglecting higher order terms, one gets

$$\dot{L} = \dot{\nu} \Rightarrow \frac{1}{\dot{L}} = \frac{r^2}{h}$$

By making unse of the following relations

$$b = \sqrt{pa}, \quad n = h/ab, \quad \dot{\nu} = h/r^2$$

the equations can be rewritten as

$$\begin{aligned}
\frac{da}{dL} &= \frac{2r^2a^2}{h^2}\left[(P_2 \sin L - P_1 \cos L)\,\varepsilon \cos\beta \cos\alpha + \frac{p}{r}\varepsilon \cos\beta \sin\alpha\right] \\
\frac{dP_1}{dL} &= \frac{r^3}{h^2}\left\{-\frac{p}{r}\cos L \varepsilon \cos\beta \cos\alpha + \left[P_1 + \left(1+\frac{p}{r}\right)\sin L\right]\varepsilon \cos\beta \sin\alpha\right. \\
&- \left. P_2\,(Q_1 \cos L - Q_2 \sin L)\,\varepsilon \sin\alpha\right\} \\
\frac{dP_2}{dL} &= \frac{r^3}{h^2}\left\{\frac{p}{r}\sin L \varepsilon \cos\beta \cos\alpha + \left[P_2 + \left(1+\frac{p}{r}\right)\cos L\right]\varepsilon \cos\beta \sin\alpha\right. \\
&+ \left. P_1\,(Q_1 \cos L - Q_2 \sin L)\,\varepsilon \sin\beta\right\} \\
\frac{dQ_1}{dL} &= \frac{r^3}{2h^2}\left(1+Q_1^2+Q_2^2\right)\sin L \varepsilon \sin\beta \\
\frac{dQ_2}{dL} &= \frac{r^3}{2h^2}\left(1+Q_1^2+Q_2^2\right)\cos L \varepsilon \sin\beta \\
\frac{dl}{dL} &= \frac{r^2 n}{h} - \frac{r^3}{h^2}\left\{\left[\frac{a}{a+b}\left(\frac{p}{r}\right)(P_1 \sin L + P_2 \cos L) + \right.\right. \\
&+ \left.\left.\frac{2b}{a}\right]\varepsilon \cos\beta \cos\alpha + \frac{a}{a+b}\left(1+\frac{p}{r}\right)(P_1 \cos L - P_2 \sin L)\,\varepsilon \cos\beta \sin\alpha + \right. \\
&+ \left. (Q_1 \cos L - Q_2 \sin L)\,\varepsilon \sin\beta\right\}
\end{aligned}$$

Introducing now the first order perturbative expansions for all the relevant quantities, the equinoctial elements are expressed as

$$\begin{aligned}
a &= a_0 + \varepsilon a_1 \\
P_1 &= P_{10} + \varepsilon P_{11} \\
P_2 &= P_{20} + \varepsilon P_{21} \\
Q_1 &= Q_{10} + \varepsilon Q_{11} \\
Q_2 &= Q_{20} + \varepsilon Q_{21} \\
l &= l_0 + \varepsilon l_1
\end{aligned}$$

Given the following variables

$$p_0 = a_0 \left[1 - (P_{10}^2 + P_{20}^2)\right] \quad h_0 = \sqrt{p_0}$$

substituting these quantities in the previous system it is possible to express those terms that give rise to first–order perturbative terms, provided that the following quantities are introduced:

$$\frac{r}{h} = \frac{h}{\mu \left(1 + P_1 \sin L + P_2 \cos L\right)}, \quad \frac{p}{r} = 1 + P_1 \sin L + P_2 \cos L$$

Collecting the terms, one obtains the following equations:

$$
\begin{aligned}
a_0' + \varepsilon a_1' &= \varepsilon \left\{ 2h_0^2 a_0^2 \cos\beta \cos\alpha \left[P_{20} \left(\frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} \right) + \right.\right. \\
&- \left.\left. P_{10} \left(\frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} \right) \right] \right. \\
&+ \left. 2h_0^2 a_0^2 \cos\beta \sin\alpha \left[\frac{1}{(1 + P_{10}\sin L + P_{20}\cos L)} \right] \right\} \\
P_{10}' + \varepsilon P_{11}' &= \varepsilon \left\{ -h_0^4 \cos\beta \cos\alpha \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} + \right. \\
&+ h_0^4 P_{10} \cos\beta \sin\alpha \frac{1}{(1 + P_{10}\sin L + P_{20}\cos L)^3} + \\
&+ h_0^4 \cos\beta \sin\alpha \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} + \\
&+ h_0^4 \cos\beta \sin\alpha \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} + \\
&- h_0^4 P_{20} Q_{10} \sin\beta \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} + \\
&+ \left. h_0^4 P_{20} Q_{20} \sin\beta \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} \right\} \\
P_{20}' + \varepsilon P_{21}' &= \varepsilon \left\{ h_0^4 \cos\beta \cos\alpha \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} + \right. \\
&+ h_0^4 P_{20} \cos\beta \sin\alpha \frac{1}{(1 + P_{10}\sin L + P_{20}\cos L)^3} + \\
&+ h_0^4 \cos\beta \sin\alpha \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} \\
&+ h_0^4 \cos\beta \sin\alpha \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^2} + \\
&+ h_0^4 P_{10} Q_{10} \sin\beta \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} + \\
&- \left. h_0^4 P_{10} Q_{20} \sin\beta \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} \right\} \\
Q_{10}' + \varepsilon Q_{11}' &= \varepsilon \left\{ \frac{h_0^4}{2} \left(1 + Q_1^2 + Q_2^2\right) \sin\beta \frac{\sin L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} \right\} \\
Q_{20}' + \varepsilon Q_{21}' &= \varepsilon \left\{ \frac{h_0^4}{2} \left(1 + Q_1^2 + Q_2^2\right) \sin\beta \frac{\cos L}{(1 + P_{10}\sin L + P_{20}\cos L)^3} \right\}
\end{aligned}
$$

where the symbol ' denotes the derivation with respect to L.

Again, for the last term there are contributions to the first–order term

from the zero–order one when substituting the perturbative expansions into it. In particular it becomes

$$\frac{\left(1-\left(\left(P_{10}+\varepsilon P_{11}\right)^2+\left(P_{20}+\varepsilon P_{21}\right)^2\right)\right)^{3/2}}{\left(1+\left(P_{10}+\varepsilon P_{11}\right)\sin L+\left(P_{20}+\varepsilon P_{21}\right)\cos L\right)^2}\Rightarrow$$

$$\Rightarrow-\sqrt{1-P_{10}^2-P_{20}^2}\left[\left(3P_{11}P_{10}P_{20}-P_{21}\left(2P_{10}^2-P_{20}^2-2\right)\right)\times\right.$$

$$\times\frac{\cos L}{\left(1+P_{10}\sin L+P_{20}\cos L\right)^3}+3\left(P_{10}P_{11}+P_{21}P_{20}\right)\frac{1}{\left(1+P_{10}\sin L+P_{20}\cos L\right)^3}$$

$$\left.+\left(P_{10}\left(P_{10}^2-2\left(P_{20}^2-1\right)\right)+3P_{10}P_{20}P_{21}\right)\frac{\sin L}{\left(1+P_{10}\sin L+P_{20}\cos L\right)^3}\right]=\mathcal{O}\left(\varepsilon\right)$$

so that

$$
\begin{aligned}
l_0' + \varepsilon l_1' \;=\;& \frac{h_0^4}{a\sqrt{a_0 p_0}} \frac{1}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^2} + \\
+\;& \varepsilon \left\{ -\frac{h_0^4 a_0 \cos\beta\cos\alpha}{a_0 + \sqrt{a_0 p_0}} \left(P_{10} \frac{\sin L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^2} + \right.\right. \\
+\;& \left.\left. P_{20} \frac{\cos L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^2} \right) + \right. \\
-\;& \frac{h_0^4 2\sqrt{a_0 p_0}\cos\beta\cos\alpha}{a_0} \frac{1}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} + \\
-\;& \frac{h_0^4 a_0 \cos\beta\sin\alpha}{a_0 + \sqrt{a_0 p_0}} \left(P_{10} \frac{\cos L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} + \right. \\
-\;& \left. P_{20} \frac{\sin L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} \right) + \\
-\;& \frac{h_0^4 a_0 \cos\beta\sin\alpha}{a_0 + \sqrt{a_0 p_0}} \left(P_{10} \frac{\cos L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^2} + \right. \\
-\;& \left. P_{20} \frac{\sin L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^2} \right) + \\
-\;& h_0^4 \sin\beta \left(Q_{10} \frac{\cos L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} + \right. \\
-\;& \left. Q_{20} \frac{\sin L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} \right) + \\
-\;& \sqrt{1 - P_{10}^2 - P_{20}^2} \left[\left(3 P_{11} P_{10} P_{20} - P_{21}\left(2P_{10}^2 - P_{20}^2 - 2\right)\right) \times \right. \\
\times\;& \frac{\cos L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} + \\
+\;& 3\left(P_{10}P_{11} + P_{21}P_{20}\right) \frac{1}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} \\
+\;& \left. \left. \left(P_{10}\left(P_{10}^2 - 2\left(P_{20}^2 - 1\right)\right) + 3 P_{10} P_{20} P_{21}\right) \frac{\sin L}{\left(1 + P_{10}\sin L + P_{20}\cos L\right)^3} \right] \right\}
\end{aligned}
$$

One can note that once L is known the information given by the last quantity l is redudant, i.e. not stricly necessary for defining vehicle position. Exaclty as in the previous case, when perturbative expansion was applied to classic orbit parameters, there is no loss of information only if time is disregarded. Another equation is needed to close the system, if time is to be evaluated. From the equation used to perform the change in derivation by means of the chain rule

$$\dot{L} = \dot{\nu} \Rightarrow \frac{1}{\dot{L}} = \frac{r^2}{h}$$

thus one can obtain a differential equation for time as a function of L:

$$\frac{dt}{dL} = \frac{r^2}{h} = \frac{h^3}{\mu \left(1 + P_1 \sin L + P_2 \cos L\right)^2} \tag{2.6}$$

which is similar to the other ones so that it can be approximated by means of a perturbative expansion, where

$$t = t_0 + \varepsilon t_1$$

Upon substitution of the expansion into Eq. (2.6) and collecting only the zero and first–order terms, one obtains

$$\begin{aligned}
t_0' + \varepsilon t_1' &= h_0^3 \frac{1}{\left(1 + P_{10} \sin L + P_{20} \cos L\right)^2} \\
&+ \varepsilon \left\{ \frac{2a_0 \left[2P_{10}^2 P_{21} - 3P_{10}P_{11}P_{20} - P_{21}\left(P_{20}^2 + 2\right)\right] - 3a_1 P_{20}\left(P_{10}^2 + P_{20}^2 - 1\right)}{2} \times \right. \\
&\times \frac{\cos L}{\left(1 + P_{10} \sin L + P_{20} \cos L\right)^3} + \\
&- \frac{2a_0 \left[P_{10}^2 P_{11} + 3P_{10}P_{20}P_{21} + 2P_{11}\left(1 - P_{20}^2\right)\right] + 3a_{11} P_{10}\left(P_{10}^2 + P_{20}^2 - 1\right)}{2} \times \\
&\times \frac{\sin L}{\left(1 + P_{10} \sin L + P_{20} \cos L\right)^3} + \\
&- 3\frac{2a_0\left(P_{10}P_{11} + P_{20}P_{21}\right) + a_{11}\left(P_{10}^2 + P_{20}^2 - 1\right)}{2} \times \\
&\times \left. \frac{1}{\left(1 + P_{10} \sin L + P_{20} \cos L\right)^3} \right\} \sqrt{a_0\left(1 - P_{10}^2 - P_{20}^2\right)}
\end{aligned}$$

Solving for the zero–order terms one gets

$$\begin{aligned}
a_0 &= \bar{a}_0 \\
P_{10} &= \bar{P}_{10} \\
P_{20} &= \bar{P}_{20} \\
Q_{10} &= \bar{Q}_{10} \\
Q_{20} &= \bar{Q}_{20} \\
l_0 &= \bar{l}_0 + \frac{\bar{h}_0^4}{\bar{a}_0 \sqrt{\bar{a}_0 \bar{p}_0}} \, I_{12}|_{L_0}^{L} \\
t_0 &= h_0^3 \, I_{12}|_{L_0}^{L}
\end{aligned}$$

where the values with an overbar are referred to quantities assigned at the initial time t_0. Solving for the terms of order one one gets

$$
\begin{aligned}
a_1 &= 2h_0^2 a_0^2 \cos\beta\cos\alpha \left[P_{20}\ I_{s2}|_{L_0}^{L} - P_{10}\ I_{c2}|_{L_0}^{L}\right] + 2h_0^2 a_0^2 \cos\beta\sin\alpha\ I_{11}|_{L_0}^{L} \\
P_{11} &= -h_0^4 \cos\beta\cos\alpha\ I_{c2}|_{L_0}^{L} + h_0^4 P_{10}\cos\beta\sin\alpha\ I_{13}|_{L_0}^{L} + h_0^4\cos\beta\sin\alpha\ I_{s3}|_{L_0}^{L} + \\
&+ h_0^4\cos\beta\sin\alpha\ I_{s2}|_{L_0}^{L} - h_0^4 P_{20}Q_{10}\sin\beta\ I_{c3}|_{L_0}^{L} + h_0^4 P_{20}Q_{20}\sin\beta\ I_{s3}|_{L_0}^{L} \\
P_{21} &= h_0^4\cos\beta\cos\alpha\ I_{s2}|_{L_0}^{L} + h_0^4 P_{20}\cos\beta\sin\alpha\ I_{13}|_{L_0}^{L} + h_0^4\cos\beta\sin\alpha\ I_{c3}|_{L_0}^{L} + \\
&+ h_0^4\cos\beta\sin\alpha\ I_{c2}|_{L_0}^{L} + h_0^4 P_{10}Q_{10}\sin\beta\ I_{c3}|_{L_0}^{L} - h_0^4 P_{10}Q_{20}\sin\beta\ I_{s3}|_{L_0}^{L}
\end{aligned}
$$

$$
\begin{aligned}
Q_{11} &= \tfrac{h_0^4}{2}\left(1+Q_1^2+Q_2^2\right)\sin\beta\ I_{s3}|_{L_0}^{L} \\
Q_{21} &= \tfrac{h_0^4}{2}\left(1+Q_1^2+Q_2^2\right)\sin\beta\ I_{c3}|_{L_0}^{L} \\
l_1 &= -\frac{h_0^4 a_0\cos\beta\cos\alpha}{a_0+\sqrt{a_0 p_0}}\left(P_{10}\ I_{s2}|_{L_0}^{L} + P_{20}\ I_{c2}|_{L_0}^{L}\right) + \\
&- h_0^4\sin\beta\left(Q_{10}\ I_{c3}|_{L_0}^{L} - Q_{20}\ I_{s3}|_{L_0}^{L}\right) \\
&- \frac{h_0^4 a_0\cos\beta\sin\alpha}{a_0+\sqrt{a_0 p_0}}\left(P_{10}\ I_{c3}|_{L_0}^{L} - P_{20}\ I_{s3}|_{L_0}^{L}\right) + \\
&- \frac{h_0^4 a_0\cos\beta\sin\alpha}{a_0+\sqrt{a_0 p_0}}\left(P_{10}\ I_{c2}|_{L_0}^{L} - P_{20}\ I_{s2}|_{L_0}^{L}\right) + \\
&- \frac{h_0^4 2\sqrt{a_0 p_0}\cos\beta\cos\alpha}{a_0}\ I_{13}|_{L_0}^{L} - \sqrt{1-P_{10}^2-P_{20}^2}\left[3\left(P_{10}P_{11}+P_{21}P_{20}\right)\ I_{13}|_{L_0}^{L} + \right. \\
&+ \left(3P_{11}P_{10}P_{20} - P_{21}\left(2P_{10}^2 - P_{20}^2 - 2\right)\right)\ I_{c3}|_{L_0}^{L} + \\
&+ \left.\left(P_{10}\left(P_{10}^2 - 2\left(P_{20}^2-1\right)\right) + 3P_{10}P_{20}P_{21}\right)\ I_{s3}|_{L_0}^{L}\right] \\
t_1 &= \left\{\frac{2a_0\left[2P_{10}^2P_{21} - 3P_{10}P_{11}P_{20} - P_{21}\left(P_{20}^2+2\right)\right] - 3a_1P_{20}\left(P_{10}^2+P_{20}^2-1\right)}{2}\ I_{c3}|_{L_0}^{L} + \right. \\
&- \frac{2a_0\left[P_{10}^2P_{11} + 3P_{10}P_{20}P_{21} + 2P_{11}\left(1-P_{20}^2\right)\right] + 3a_{11}P_{10}\left(P_{10}^2+P_{20}^2-1\right)}{2}\ I_{s3}|_{L_0}^{L} + \\
&- \left.3\frac{2a_0\left(P_{10}P_{11}+P_{20}P_{21}\right) + a_{11}\left(P_{10}^2+P_{20}^2-1\right)}{2}\ I_{13}|_{L_0}^{L}\right\}\sqrt{a_0\left(1-P_{10}^2-P_{20}^2\right)}
\end{aligned}
$$

Obviously the classical orbital parameters can be calculated with the follow-

ing relations [2]

$$\begin{aligned} e^2 &= P_1^2 + P_2^2 \qquad \tan^2\frac{i}{2} = Q_1^2 + Q_2^2 \\ \tan\varpi &= \frac{P_1}{P_2} \qquad\qquad \tan\Omega = \frac{Q_1}{Q_2} \end{aligned}$$

For the system solution it must be introduced the following integrals (with $P_{10}, P_{20} \neq 0$, otherwise the solution become trivial)

$$\begin{aligned}
I_{c3} &= \int \frac{\cos L}{(1+P_{10}\sin L+P_{20}\cos L)^3}dL = \\
&= -\frac{3P_{20}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\Big]}{2\Big(P_{10}^2+P_{20}^2-1\Big)^{5/2}}+ \\
&+ \frac{3P_{20}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\Big]}{2\Big(P_{10}^2+P_{20}^2-1\Big)^{5/2}}+ \\
&+ \frac{P_{10}\Big(P_{10}^4+P_{10}^2P_{20}\Big(2P_{20}^2+P_{20}-2\Big)+P_{20}\Big(P_{20}^4+4P_{20}^3-5P_{20}^2-P_{20}+1\Big)\Big)\cos^2 L}{2P_{20}\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{P_{20}^2\cos L\Big(\Big(2P_{10}^2+2P_{10}^2\Big(2P_{20}^2+P_{20}-3\Big)+2P_{20}^4-4P_{20}^3+3P_{20}^2-2P_{20}+1\Big)\sin L+P_{10}\Big(2P_{10}^2+5P_{20}^2-5\Big)\Big)}{2P_{20}\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{P_{10}\Big(P_{10}^4(P_{20}+1)+4P_{10}^2P_{20}\Big(P_{20}^2-1\Big)+P_{20}\Big(2P_{20}^4-3P_{20}^3+2P_{20}^2-3P_{20}+2\Big)\Big)\sin^2 L}{2P_{20}\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{P_{20}\Big(2P_{10}^4+2P_{10}^2\Big(3P_{20}^2-P_{20}-2\Big)+P_{20}^4-2P_{20}^3+3P_{20}^2-4P_{20}+2\Big)\sin L}{2P_{20}\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{P_{10}\Big(P_{10}^4+P_{10}^2P_{20}\Big(2P_{20}^2-P_{20}-2\Big)+P_{20}\Big(P_{20}^4-P_{20}^3-5P_{20}^2+4P_{20}+1\Big)\Big)}{2P_{20}\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2} \\
I_{c2} &= \int \frac{\cos L}{(1+P_{10}\sin L+P_{20}\cos L)^3}dL = \\
&= \frac{P_{20}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big]}{\Big(P_{10}^2+(P_{20}+1)(P_{20}-1)\Big)^{3/2}}+ \\
&- \frac{P_{20}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big]}{\Big(P_{10}^2+(P_{20}+1)(P_{20}-1)\Big)^{3/2}}+ \\
&+ \frac{P_{10}P_{20}\cos L+\Big(P_{10}^2+P_{20}-1\Big)\sin L+P_{10}P_{20}}{(1-P_{20})\Big(P_{10}^2+P_{20}^2-1\Big)(P_{20}\cos L+P_{10}\sin L+1)}
\end{aligned}$$

$$
\begin{aligned}
I_{s3} &= \int \frac{\sin L}{(1+P_{10}\sin L+P_{20}\cos L)^3}dL = \\
&= \frac{3P_{10}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\sqrt{P_{10}^2+P_{20}^2}+P_{10}\Big]}{2\Big(P_{10}^2+P_{20}^2-1\Big)^{5/2}}+ \\
&- \frac{3P_{10}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\sqrt{P_{10}^2+P_{20}^2}-P_{10}\Big]}{2\Big(P_{10}^2+P_{20}^2-1\Big)^{5/2}}+ \\
&+ \frac{\Big(P_{10}^2\Big(2P_{20}^2-2P_{20}-1\Big)+P_{20}\Big(P_{20}^3+P_{20}^2-P_{20}-1\Big)\Big)\cos^2 L}{2(1-P_{20})\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{\cos L\Big(P_{10}\Big(2P_{10}^2-4P_{20}^2+3P_{20}+1\Big)\sin L+\Big(P_{10}^2-2\Big(P_{20}^2-1\Big)\Big)(P_{20}+1)\Big)}{2(1-P_{20})\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{P_{10}^2\Big(3(P_{20}-1)-P^2\Big)\sin^2 L-P_{10}\Big(2P_{10}^2-P_{20}^2-3P_{20}+4\Big)\sin L}{2(1-P_{20})\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&+ \frac{P_{10}^2P_{20}(1-2P_{20})-P_{20}^4+P_{20}^3+3P_{20}^2-P_{20}-2}{2(1-P_{20})\Big(P_{10}^2+P_{20}^2-1\Big)^2(P_{20}\cos L+P_{10}\sin L+1)^2} \\
I_{s2} &= \int \frac{\sin L}{(1+P_{10}\sin L+P_{20}\cos L)^3}dL = \\
&= \frac{P_{10}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}-P\Big]}{\Big(P_{10}^2+P_{20}^2-1\Big)^{3/2}}+ \\
&- \frac{P_{10}\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big]}{\Big(P_{10}^2+P_{20}^2-1\Big)^{3/2}}+ \\
&+ \frac{(P_{20}+1)\cos L+P_{10}\sin L+P_{20}+1}{\Big(P_{10}^2+P_{20}^2-1\Big)(P_{20}\cos L+P_{10}\sin L+1)} \\
I_{11} &= \int \frac{1}{(1+P_{10}\sin L+P_{20}\cos L)}dL = \\
&= \frac{\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big]}{\sqrt{P_{10}^2+P_{20}^2-1}}+ \\
&- \frac{\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big]}{\sqrt{P_{10}^2+P_{20}^2-1}} \\
I_{12} &= \int \frac{1}{(1+P_{10}\sin L+P_{20}\cos L)^2}dL = \\
&= \frac{\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big)\cos L+(P_{20}-1)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\Big]}{\Big(P_{10}^2+P_{20}^2-1\Big)^{3/2}}+ \\
&- \frac{\ln\Big[\Big(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big)\cos L+(1-P_{20})\sin L+\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\Big]}{\Big(P_{10}^2+P_{20}^2-1\Big)^{3/2}}+ \\
&+ \frac{P_{10}\cos L+\Big(P_{10}^2+P_{20}(P_{20}-1)\Big)\sin L+P_{10}}{\Big(P_{10}^2+P_{20}^2-1\Big)(P_{20}-1)(1+P_{10}\sin L+P_{20}\cos L)}
\end{aligned}
$$

$$
\begin{aligned}
I_{13} &= \int \frac{1}{(1+P_{10}\sin L+P_{20}\cos L)^3}dL = \\
&= \frac{\left(P_{10}^2+P_{20}^2+2\right)\ln\left[\left(\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\right)\cos L+\left(P_{20}^2-1\right)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}-P_{10}\right]}{2\left(P_{10}^2+P_{20}^2-1\right)^{5/2}}+ \\
&- \frac{\left(P_{10}^2+P_{20}^2+2\right)\ln\left[\left(\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\right)\cos L+\left(1-P_{20}^2\right)\sin L+\sqrt{P_{10}^2+P_{20}^2-1}+P_{10}\right]}{2\left(P_{10}^2+P_{20}^2-1\right)^{5/2}}+ \\
&- \frac{P_{10}\left(P_{10}^4+P_{10}^2\left(3P_{20}^2-P_{20}-1\right)+P_{20}\left(2P_{20}^3-P_{20}^2+P_{20}-2\right)\right)\cos^2 L}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{\cos L\left(\left(2P_{10}^4P_{20}+P_{20}^2(P_{20}-1)\left(2P_{20}^2+5P_{20}+3\right)+3P_{20}^2(P_{20}-1)^2\right)\sin L\right)}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{\cos L\left(P_{10}\left(P_{10}^2\left(P_{20}^2+1\right)+P_{20}^4+3P_{20}^2-4\right)\right)}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{P_{10}\left(2P_{10}^4+P_{10}^2(P_{20}-1)(3P_{20}+5)+P_{20}(P_{20}+4)(P_{20}-1)^2\right)\sin^2 L}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{\left(2P_{10}^4+P_{10}^2(1-P_{20})\left(P_{20}^2-3P_{20}-8\right)-P_{20}(P_{20}-1)^2\left(P_{20}^2-4\right)\right)\sin L}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}+ \\
&- \frac{P_{10}\left(P_{10}^4+P^2\left(2P_{20}^2-P_{20}-2\right)+P_{20}^4-P_{20}^3-2P_{20}^2-2P_{20}+4\right)}{2\left(P_{10}^2+P_{20}^2-1\right)^2(P_{20}-1)^2(P_{20}\cos L+P_{10}\sin L+1)^2}
\end{aligned}
$$

which must be calculated each time between L and L_0.

In case one wants to express the equations for a thrust vector parallel to the velocity direction, simply note that, letting γ be the flight path angle

$$\cos\gamma = \frac{r\dot{\nu}}{v} \quad \sin\gamma = \frac{\dot{r}}{v}$$

and writing the velocity magnitude as

$$v = \sqrt{\frac{a}{2a\left(1+P_1\sin L+P_2\cos L\right)-h^2}}$$

the direction α of thrust with respect to the transverse (horizontal) axis equals the flight path angle, so that

$$\cos\gamma = \sin\alpha \quad \sin\gamma = \cos\alpha$$

Since $\mu = 1$ was assumed, when scaling times,

$$\begin{aligned}
\dot{r} &= \frac{dr}{dL}\dot{L} \\
\frac{dr}{dL} &= \frac{h^2\left(P_2 \sin L - P_1 \cos L\right)}{\left(1 + P_1 \sin L + P_2 \cos L\right)^2} \\
\dot{L} &= \dot{\nu} = \frac{\left(1 + P_1 \sin L + P_2 \cos L\right)^2}{h^3} \\
\dot{r} &= \frac{P_2 \sin L - P_1 \cos L}{h} \\
r\dot{\nu} &= \frac{h}{r} = \frac{\left(1 + P_1 \sin L + P_2 \cos L\right)^2}{h}
\end{aligned}$$

One thus obtains the following relations

$$\begin{aligned}
\cos\alpha &= \frac{\sqrt{a}\left(P_2 \sin L - P_1 \cos L\right)}{\sqrt{2a\left(1 + P_1 \sin L + P_2 \cos L\right) - h^2}} \\
\sin\alpha &= \frac{\sqrt{a}\left(1 + P_1 \sin L + P_2 \cos L\right)}{\sqrt{2a\left(1 + P_1 \sin L + P_2 \cos L\right) - h^2}}
\end{aligned}$$

that can be substituted in the previous equations.

The integrals reported in the previous page do not allow for an analitical solution, but they can be solved by means of a numerical quadrature scheme.

It must be underlined that these expansions work both forward and backward in angular position, simply introducing a minus sign in the value of L when computing the various quantities.

In the next three figures 2.8, 2.9 and 2.10 are reported three cases with variable thrust orientation. In the first two pictures where the out of plane angle β is equal to zero the value of inclination i and RAAN Ω remains constant, while in the last figure where the thrust is directed also out of the plane they start to vary. In all the figures the perturbation is compared with an accurate numerical solution, it is clear how standard perturbation is able to capture both the short and long period terms. The parameters trends are qualitatively similar, there is not a sharp change if the thrust orientation is modified.

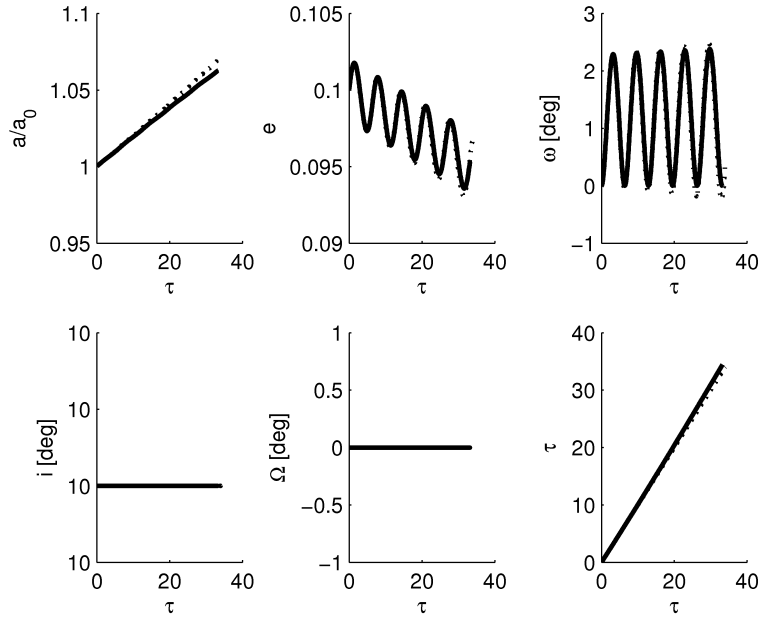

Figure 2.8: Thrust orientation: $\alpha = 90$ deg and $\beta = 0$ deg. $\varepsilon = 10^{-3}$

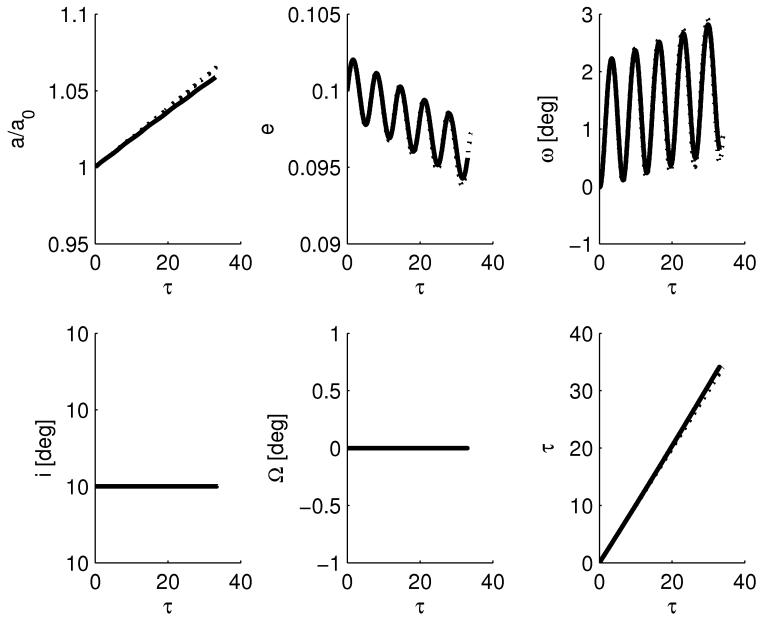

Figure 2.9: Thrust orientation: $\alpha = 70$ deg and $\beta = 0$ deg. $\varepsilon = 10^{-3}$

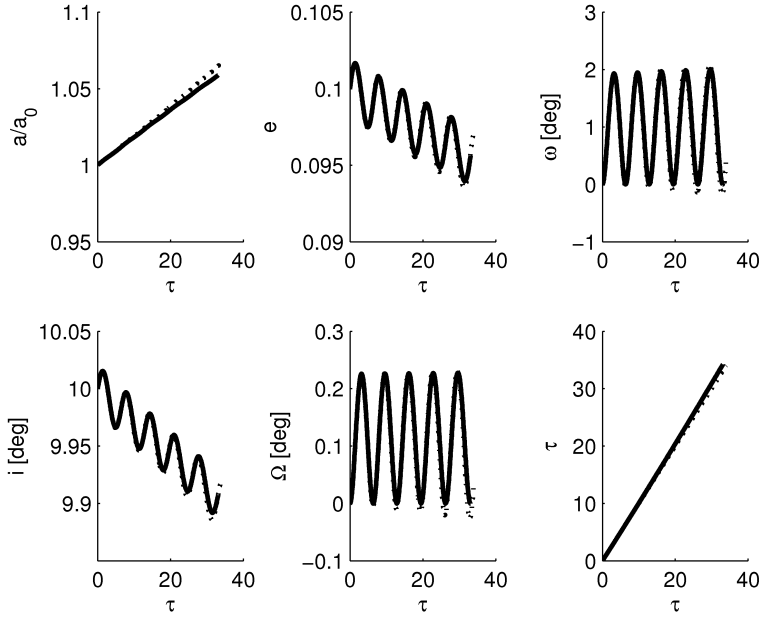

Figure 2.10: Thrust orientation: $\alpha = 90$ deg and $\beta = 20$ deg. $\varepsilon = 10^{-3}$

Chapter 3

Error Analysis for Perturbative Approximations of Low–Thrust Trajectory Arcs

In this chapter the results obtained from the perturbative expansions developed in the previous chapter are presented and critically analyzed in terms of prediction error and computational cost. In the first section the relation between thrust order of magnitude and perturbation parameter ε is analyzed. Two settings has been considered: geocentric and heliocentric ones, for different distances from the main body. In the next two sections, the approximations are compared with an accurate simulation performed by means of Encke's method and in the fourth one the accuracy in terms of perdiction error is evaluated. In the fifth section, the first order perturbative analytical solution is compared with another analytical approach for propagating orbital motion under low–thrust propulsion, based on a Fourier series expansion [18]. In the sixth section there is the computational time analysis, while , in the seventh one the use of the analytical expansions in the framework of direct optimization is considered as a possible applicative scenario. The results obtained from the approximations appear to be satisfactory provided that trajectory arc angular length, $\Delta\theta$, and values of the perturbation parameter, $\epsilon = (T/m)/(\mu/r^2)$, remain compatible, that is for higher values of ε, a smaller arc must be adopted. Taking care of these limitations, the application of the perturbative expansion to direct optimization methods appears a promising tool as the results posted in the eighth section show.

3.1 Case studies

Table 3.1 presents the list of case studies that will be considered in this chapter as a test bench-mark for both perturbative expansions developed in the previous chapter, namely, the multiple scale solution and the standard perturbation approximation. As it will be discussed later in the chapter, the error of the analytical approximations with respect to the numerical integration depends first of all on the value of the adimensionalized acceleration. This is obvious, since it is assumed that the perturbation parameter is small for finding a first order solution. In order to evaluate if the relations developed in the previous chapter are of any practical interest it is necessary to consider their accuracy for parameter values within reasonable values for practical realistic space missions. Six different settings are considered in Table 3.1, three for a geocentric scenario (LEO, GEO and near–Moon orbits) and three for a heliocentric one, at a distance of 1 AU (near Earth), of 1.5 AU (near Mars) and at a distance of 5 AU (near Jupiter). The mass of the vehicle is assumed to be of one ton. Reasonable values for the dimensionless acceleration lie between 10^{-1} and 10^{-4}, the corresponding low–thrust forces needed for obtaining these values are reported in Table 3.1 for different test cases. For the geocentric situations it is quite clear that for any realistic value of electric propulsion thrust the resulting non–dimensional acceleration would be less than 10^{-2}. For heliocentric trajectories also little values of thrust, around 10^{-2} N, can cause a dimensionless acceleration larger than 10^{-2}, if the near Jupiter case is considered. These aspects will be crucial in the analysis of the results performed in the following paragraphs.

	$\varepsilon = \frac{T/m}{\mu/r^2}$		$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-4})$
Geocentric	$R_\oplus$	T [N]	-	-	$\mathcal{O}(10)$	$\mathcal{O}(1)$
	R_{GEO}	T [N]	$\mathcal{O}(10)$	$\mathcal{O}(1)$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$
	R_{Moon}	T [N]	$\mathcal{O}(1)$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$
Heliocentric	1 AU	T [N]	$\mathcal{O}(1)$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$
	1.5AU	T [N]	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$	-
	5 AU	T [N]	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$	-	-

Case studies for a vehicle of mass $m = 1000$ kg.

Table 3.1: Thrust values for a given ϵ in different scenraios.

After a discussion focused on the variation of orbit parameters reported in the plots presented in par. 3.2 and 3.3, a detailed analysis on the error resulting from the perturbative expansion will be performed, with the aim of identifying the maximum angular distance flown along a low–thrust

trajectory that results in an acceptable error lever for a given value of the perturbation parameter, i.e. the non–dimensional acceleration.

This analysis of the error is performed assuming different values of ε and evaluating orbit parameters after a prescribed angular travel, that can be as large as 5 orbits when ε is below 10^{-3}, but is limited to few degrees for 10^{-1}. These results will then be used for evaluating practical limits for the use of the considered perturbative expansions in both the geocentric and heliocentric cases, that is the maximum distance from the primary body taht can be considered for a given set of vehicle parameters, thrust T and mass m.

3.2 Two Scale Perturbation Results

In this section the results for the two scale approach are compared to an accurate numerical solution found by means of Encke's method. There are two figures for each value of ε: one for the trend of the parameters in time and a second one reporting the variation of error with time. The formulation is two–dimensional so only three plot per figure are reported: a, e and ω. Thrust orientation is defined by means of α the angle between thrust and the trasversal direction. In figures from 3.1 to 3.8 trasversal thrust is assumed ($\alpha = 0$) and ε vary from 10^{-4} to 10^{-1}; the total travel length chosen for the first two cases was of 5 orbits, while for the other two the arc length has been chosen depending on the ε values.

For the case of trasversal thrust direction two figures rappresenting the trajectory are reported, 3.13 and 3.16, where $\epsilon = 10^-3$ and $\epsilon = 10^-2$ respectively. In addition the error variation, with respect to time, between the radius calculated by the accurate numerical solution and the one obtained by means of two scale method has been investigated for two ε values: 10^{-3} and 10^{-2} (figs. 3.14, 3.15, 3.17 and 3.18). In figures 3.9 and 3.10 a case of tangential thrust is presented and in figures 3.11 and 3.12 a thrust angle of $\alpha = 20$ deg has been used.

As it can be seen from the figures 3.1, 3.3, 3.5 and 3.7 the approximation is able to capture the long variation of the parameters, without catching the short period one. The errors for the first two cases are quite small also for a large number of revolutions. This is no longer true for the case of higher acceleration. In this situation the error remains little only for a fraction of the total trip considered. The error remains acceptable only if the expansion is used for angular travels approximately equal to one revolution (or even less).

Note that if compared with the standard perturbation results the two

scale method seem to provide a solution which is less affected by the drift from the accurate numerical one, although it does not capture the short term variations.

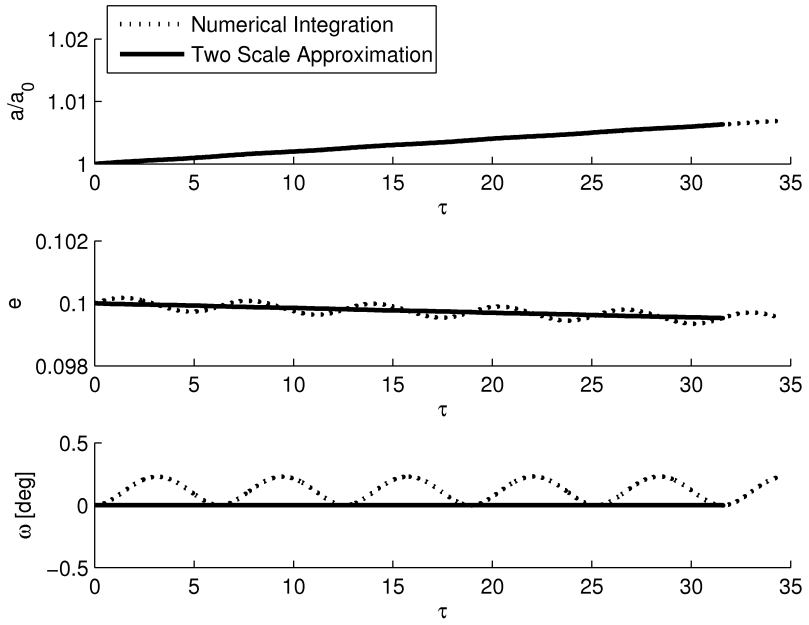

Figure 3.1: Parameters in time, $\alpha = 0$ deg, $\varepsilon = 10^{-4}$

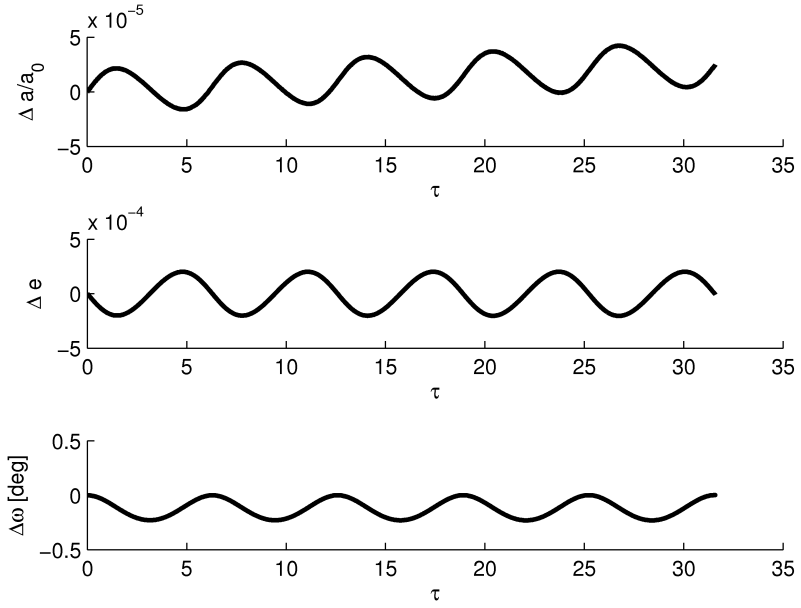

Figure 3.2: Error in time, $\alpha = 0$ deg, $\varepsilon = 10^{-4}$

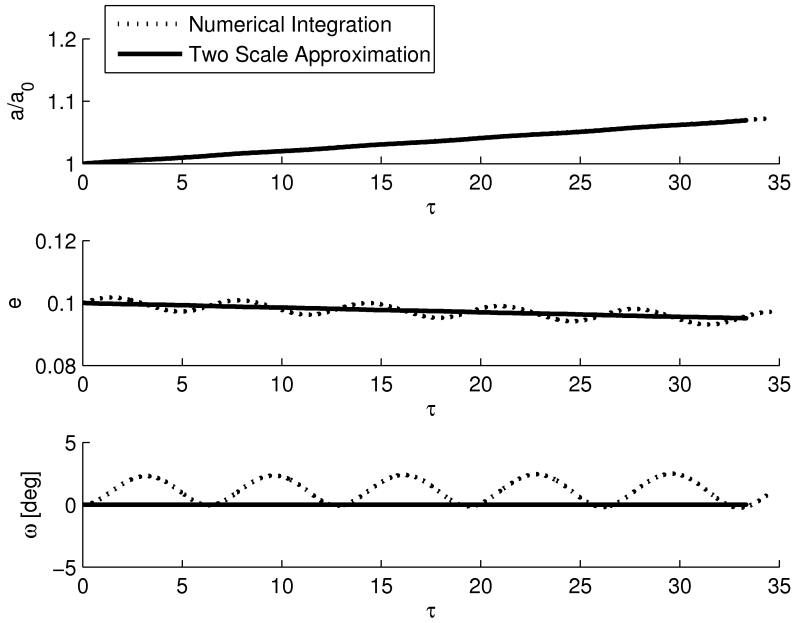

Figure 3.3: Parameters in time, $\alpha = 0$ deg, $\varepsilon = 10^{-3}$

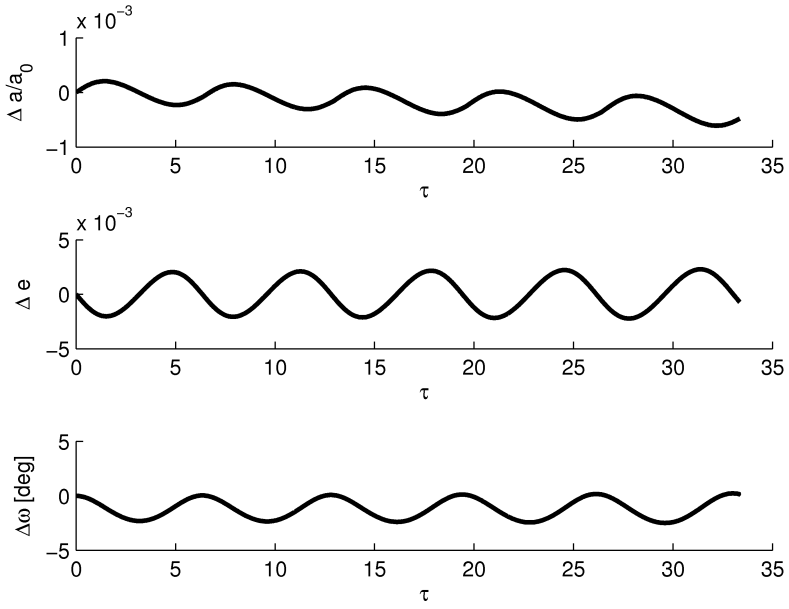

Figure 3.4: Error in time, $\alpha = 0$ deg, $\varepsilon = 10^{-3}$

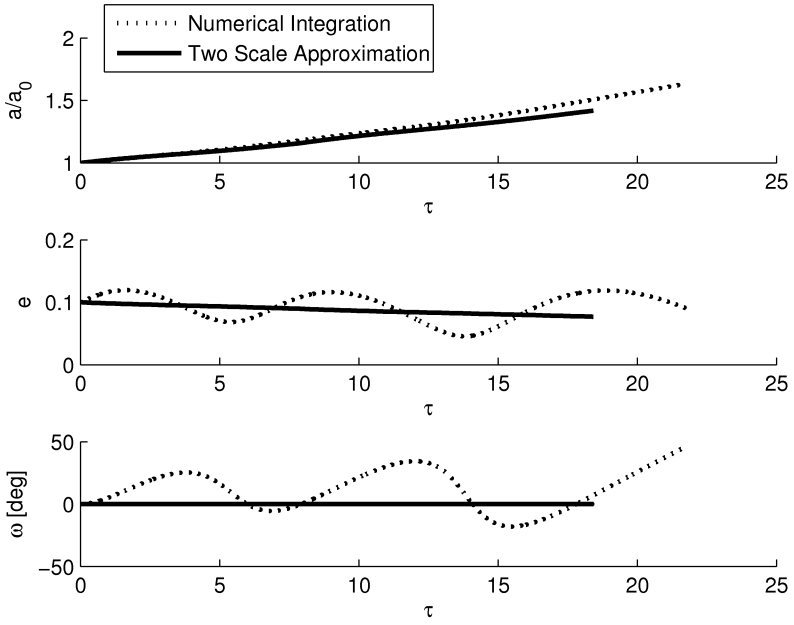

Figure 3.5: Parameters in time, $\alpha = 0$ deg, $\varepsilon = 10^{-2}$

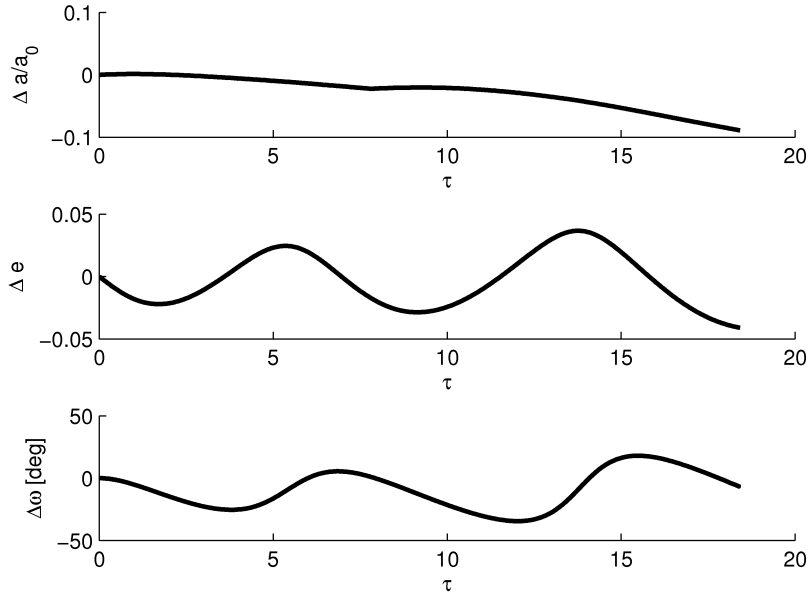

Figure 3.6: Error in time, $\alpha = 0$ deg, $\varepsilon = 10^{-2}$

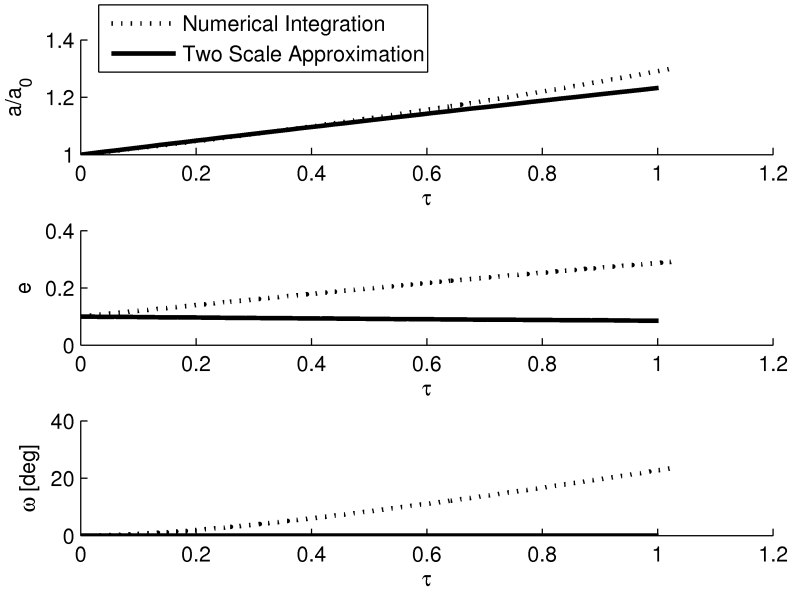

Figure 3.7: Parameters in time, $\alpha = 0$ deg, $\varepsilon = 10^{-1}$

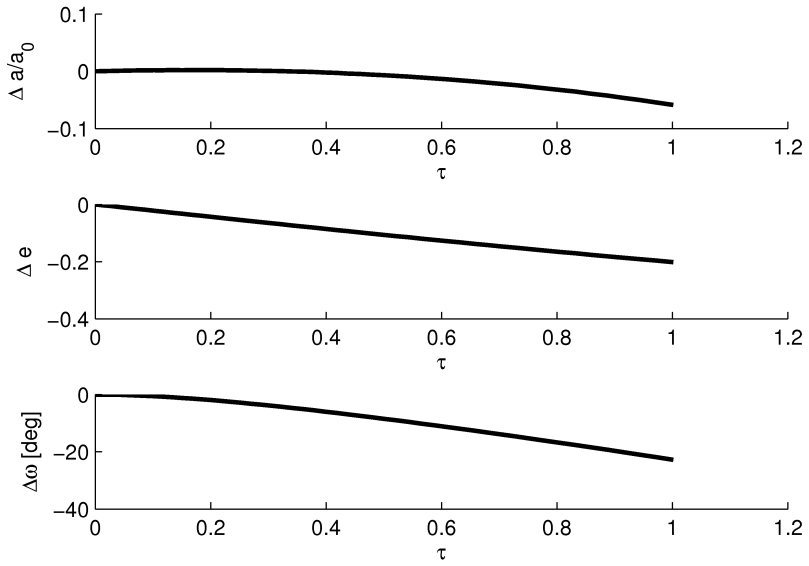

Figure 3.8: Error in time, $\alpha = 0$ deg, $\varepsilon = 10^{-1}$

From the following figures it is clear that if a tangential thrust case or a case with thrust in arbitrary direction is considered, the trend for the parameters and the error in time is qualitatively similar to the case of trasversal thrust.

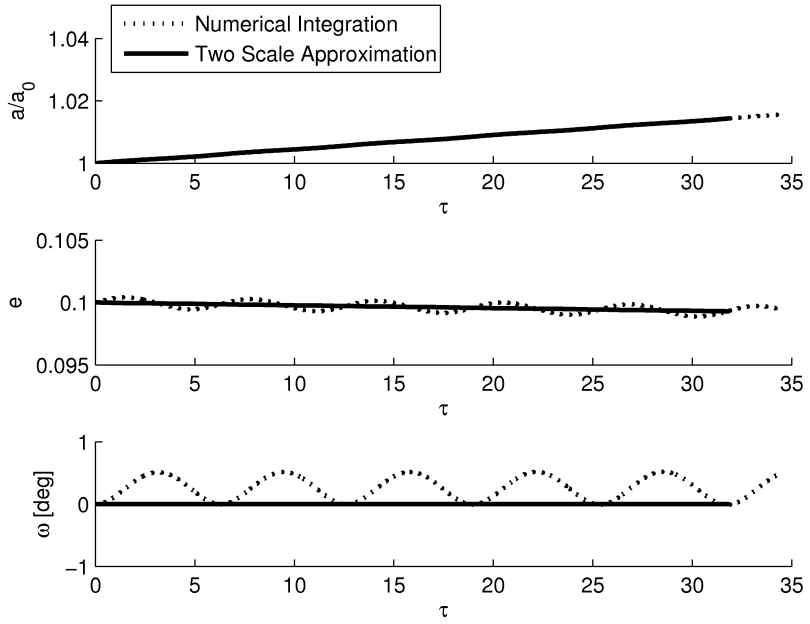

Figure 3.9: Parameters in time, $\vec{\boldsymbol{T}} \parallel \vec{\boldsymbol{V}}$, $\varepsilon = 10^{-3}$

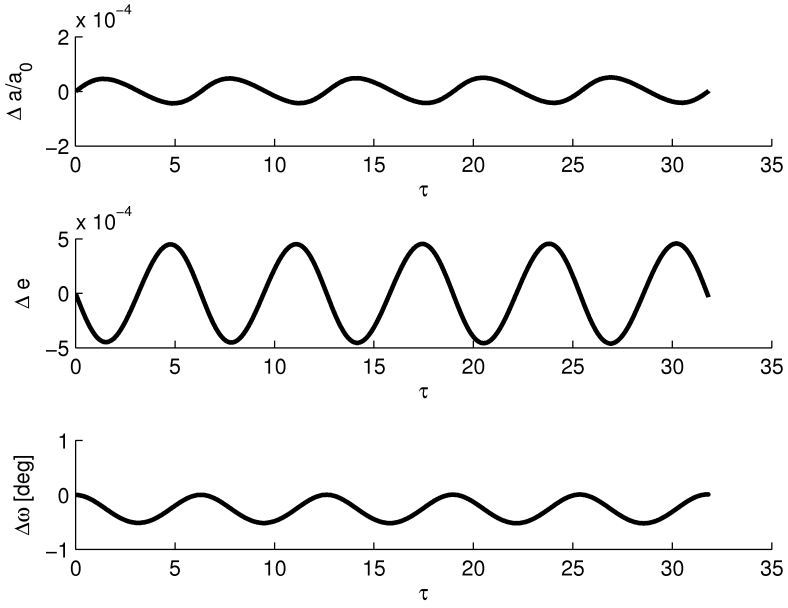

Figure 3.10: Error in time, $\vec{\boldsymbol{T}} \parallel \vec{\boldsymbol{V}}$, $\varepsilon = 10^{-3}$

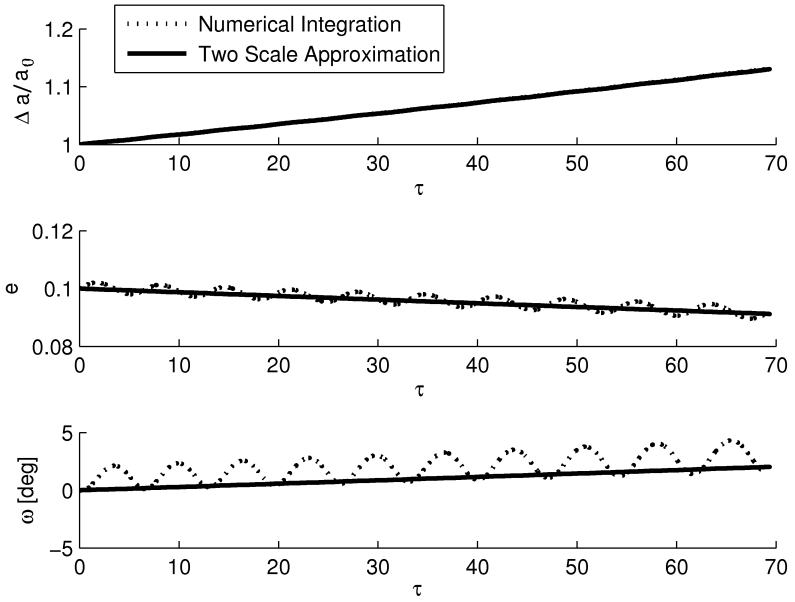

Figure 3.11: Parameters in time, $\alpha = 20$ deg, $\varepsilon = 10^{-3}$

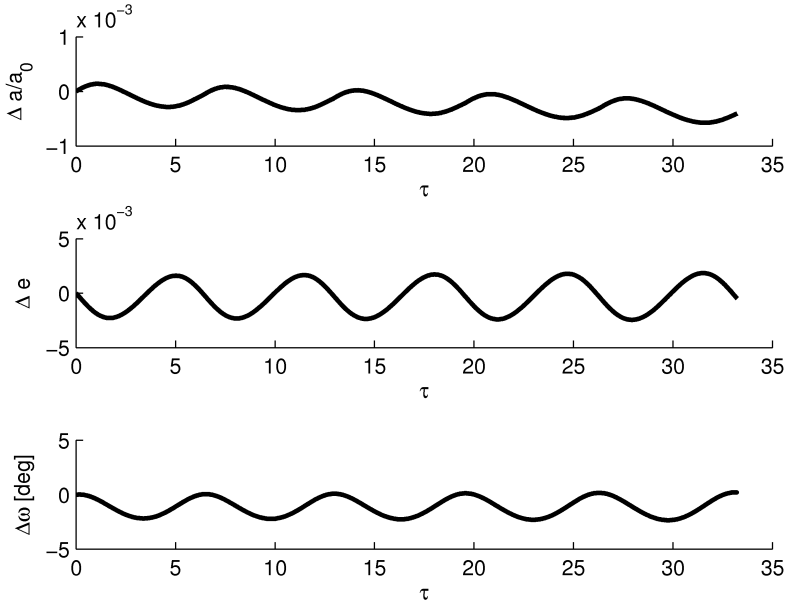

Figure 3.12: Error in time, $\alpha = 20$ deg, $\varepsilon = 10^{-3}$

Figures 3.13 and 3.16 present the trajectory shape for $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-2}$ respectively. For both cases the trajectory approximation is satisfactory: even if the two scale method is only able to capture the long period variations, it allow for obtain a good trajectory approximation (even for a large number of orbits) if ε remains sufficiently small.

In what follows the trends for the radius error, defined as[1]

$$\frac{\Delta r}{a_0} = \frac{\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}_{ts}}{a_0}$$

are presented. In 3.14 and 3.17 its value is plotted as a function of time, while in 3.15 and 3.18 it is rappresented as a function of the angular coordinate θ. It is clear that in the first two cases the error is large whereas in the second ones it remains limited to acceptable values. This fact is caused by the indipendent varible change, from time to longitude. With this action the time is become a dipendent variable so it cannot be evaluated exaclty, so that this error cause a position error in time which grows rapidly. If the position error is evaluated with respect to the angular coordinate the problem of time is overcome and the precision become higher.

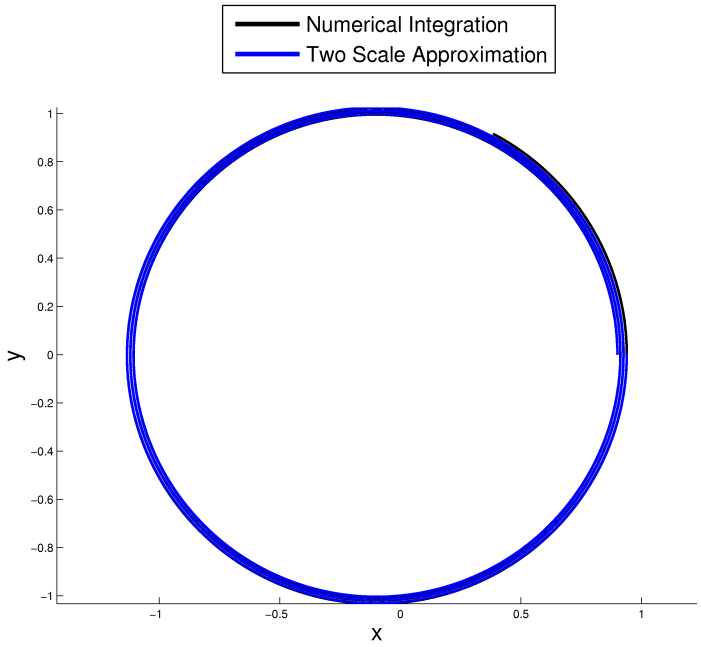

Figure 3.13: Trajectory, $\alpha = 0$ deg, $\varepsilon = 10^{-2}$

[1]The subscript ts means two scale.

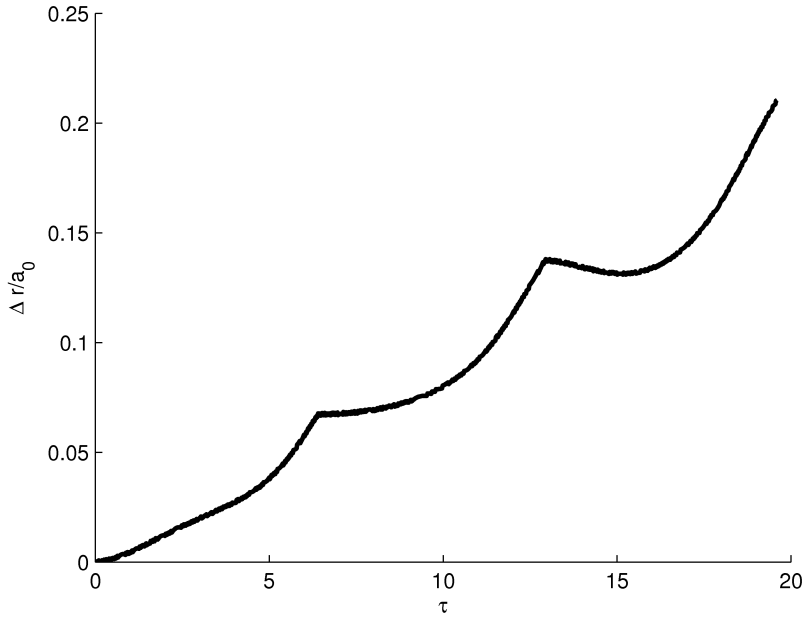

Figure 3.14: Error on the radius, $\alpha = 0$ deg, function of time, $\varepsilon = 10^{-2}$

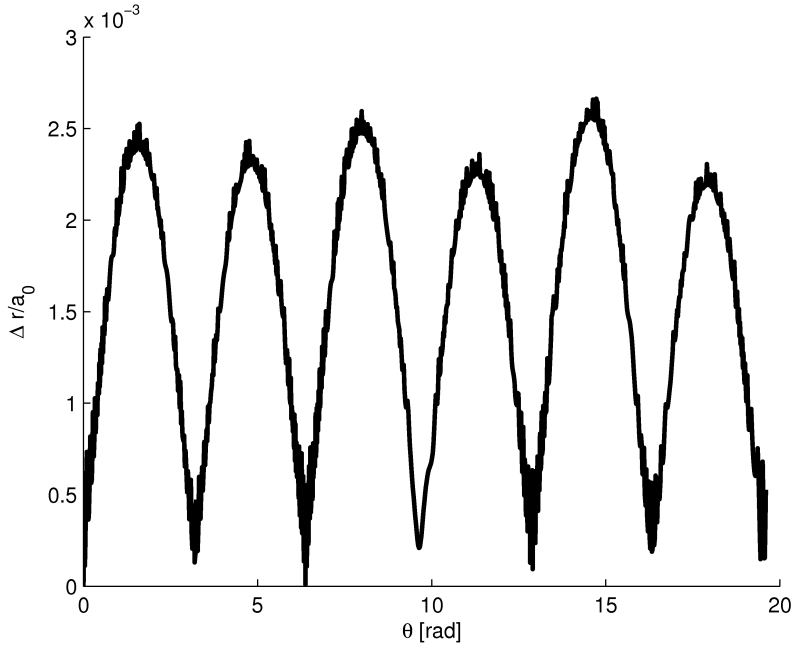

Figure 3.15: Error on the radius, $\alpha = 0$ deg, function of θ, $\varepsilon = 10^{-2}$

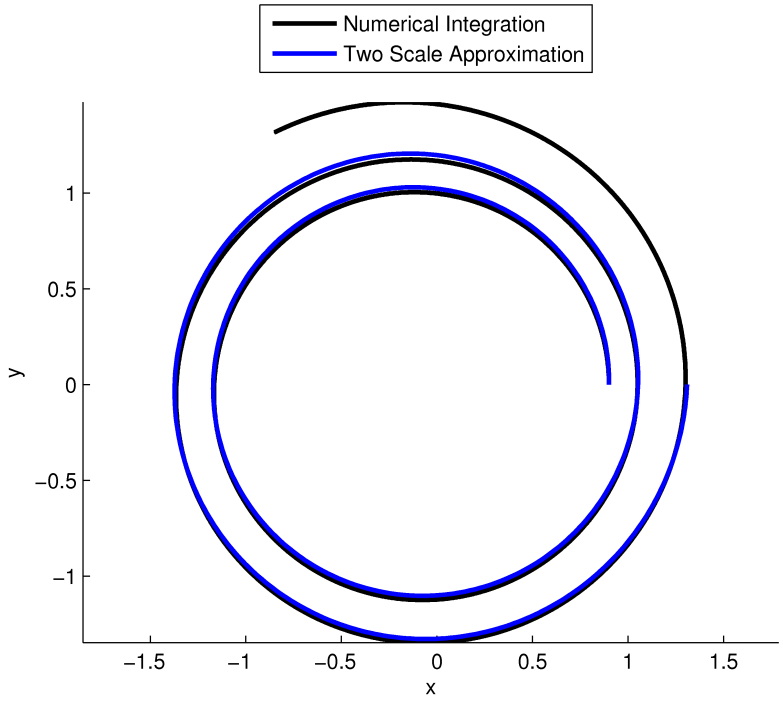

Figure 3.16: Trajectory, $\alpha = 0$ deg, $\varepsilon = 10^{-2}$

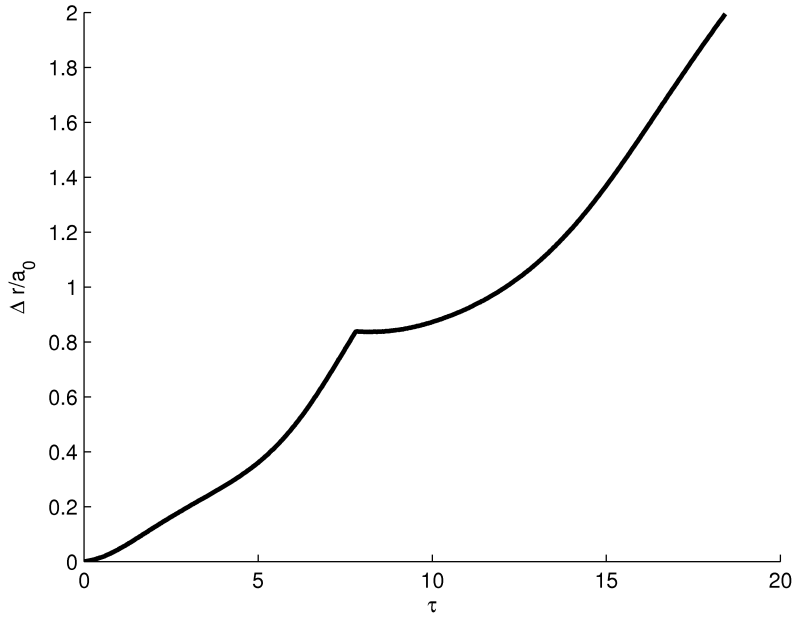

Figure 3.17: Error on the radius, $\alpha = 0$ deg, function of time, $\varepsilon = 10^{-2}$

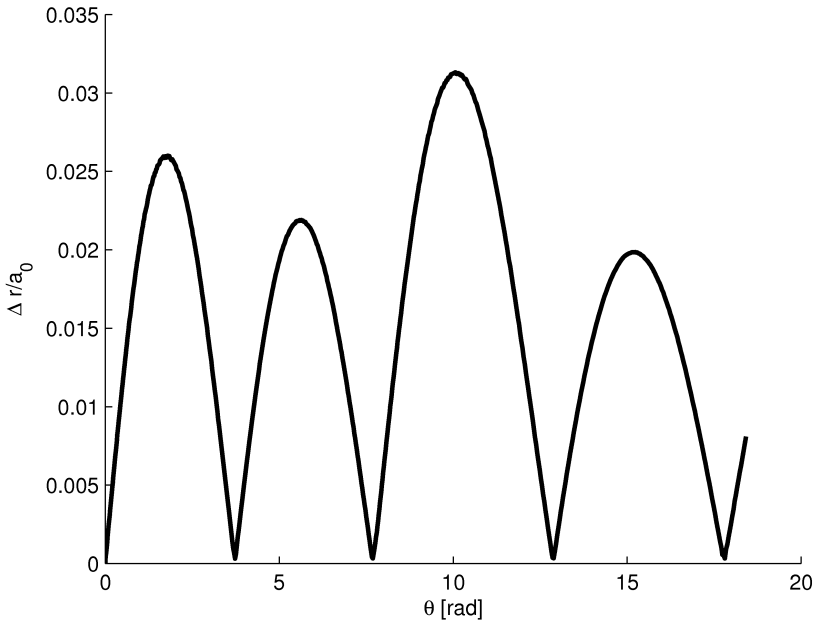

Figure 3.18: Error on the radius, $\alpha = 0$ deg, function of θ, $\varepsilon = 10^{-2}$

3.3 Standard Perturbtion Results

In this section the results for the standard perturbation approach are compared to an accurate numerical solution found again by means of Encke's method. In what follows two cases are presented, for a general case (with non–zero eccentricity and inclination) and a circular orbit with zero inclination, in order to highlight the fact that the formulation proposed in this work is not affected by singularities.

Similarly to what was presented in the previous section, there are two figures for each value of ε: one for the trend of the parameters in time and a second one reporting the variation of error with time. Since the formulation is three–dimensional, six plots per figure are presented, where the last parameter is the flight time instead of the mean longitude, as explained in detail in the previous chapter. Two angles for the thrust orientation need to be considered: α in the orbit plane and β outside of it. Their values will be specified case by case.

For the general case for the first two values of ε ten orbits have been simulated, while for the other cases a variable number of orbits has been considered, depending on the value of ε.

3.3.1 Generic Eccentric Inclined Orbit

For this case the angles have been chosen to be $\alpha = 45$ deg and $\beta = 20$ deg. As it can be seen from figures 3.19, 3.21, 3.23 and 3.25 the approximation is able to capture both the long and the short period variations of the parameters. The errors for the first two cases are quite small also for a large number of revolutions. This is no longer true for the case of higher acceleration. In this situation the error remains little only for a fraction of the total trip considered. The error remains acceptable only if the expansion is used for angular travels approximately equal to one revolution (or even less), as already seen for the two scale method.

Note that, in all the considered cases, the error grows approximately as a quadratic function of τ, with one exception for a,e and i for $\varepsilon = 10^{-4}$, where the behaviour is almost linear, but the absolute value remains very small.

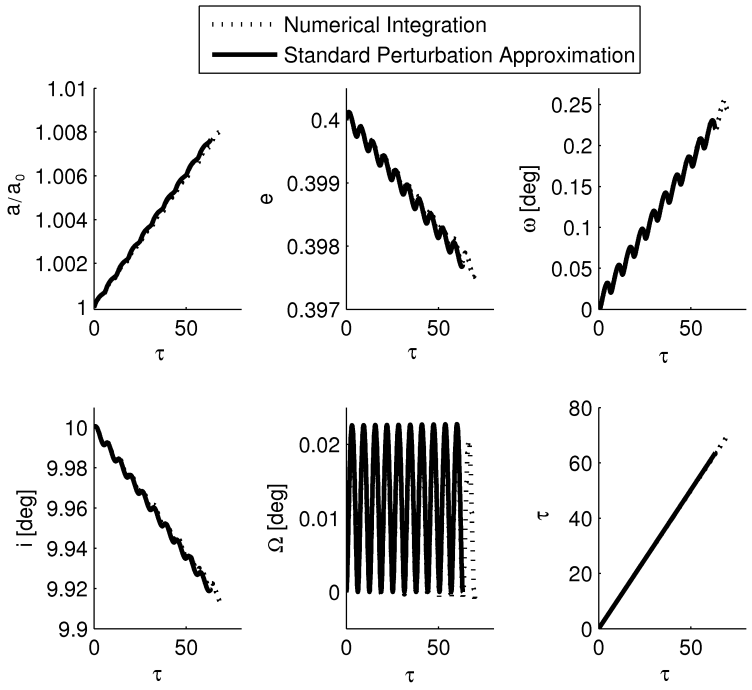

Figure 3.19: Parameters in time, $\varepsilon = 10^{-4}$

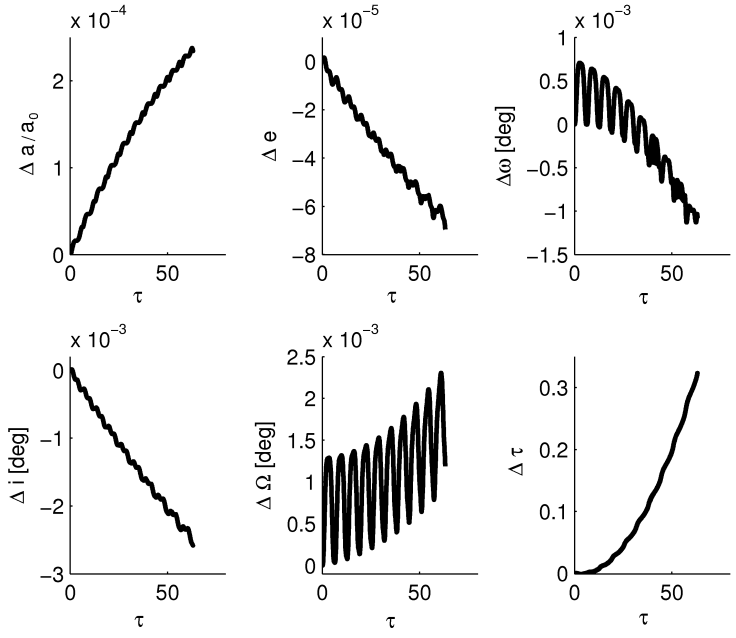

Figure 3.20: Error in time, $\varepsilon = 10^{-4}$

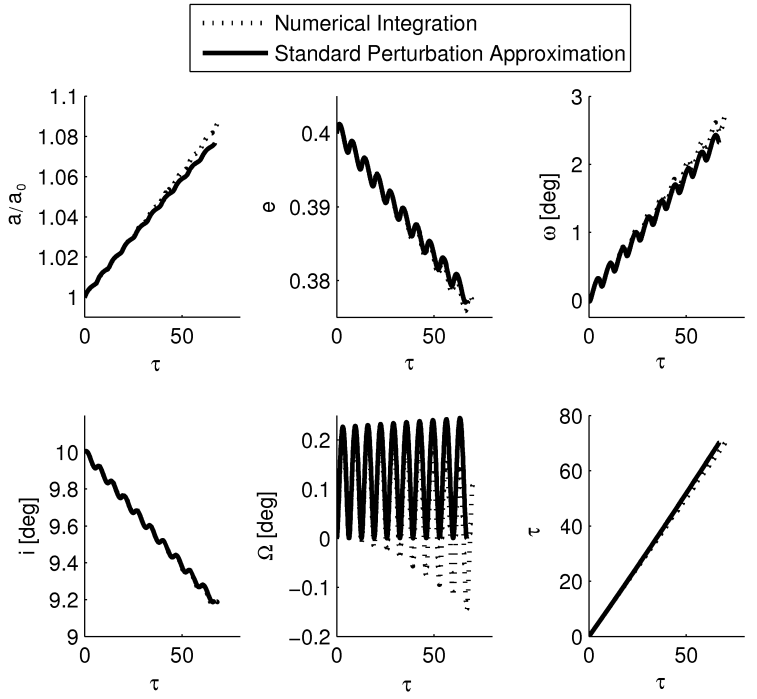

Figure 3.21: Parameters in time, $\varepsilon = 10^{-3}$

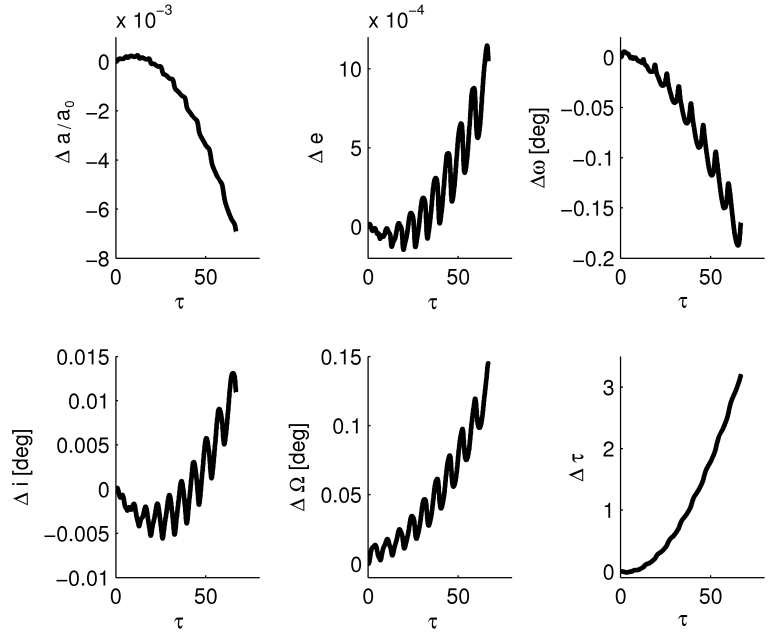

Figure 3.22: Error in time, $\varepsilon = 10^{-3}$

Another perculiar aspect of the standard perturbation approach is that is capable of capturing the short–term, suborbital oscillation of the parameters, provided that ε is sufficiently small. When ε grows, the most dramatic error is the underestimation of parameter oscillation amplitudes with respect to the actual variation estimated by the numerical propagation. For the considered cases, e, ω and Ω are particularly affected by this problem (see Figs. 3.23 and 3.24). One should note at the sime time that at this point a purely forward propagation is being considered, where the values of the integrals at the basis of the perturbative expansions are evaluated only once, for the initial conditions. A significant improvement in error performance can be obtained if these integrals are updated every once in a while, as it will be done in section 4.4, for comparing the results obtained by means of the perturbative approaches with those obtained by Scheeres & Hudson in Ref. [18].

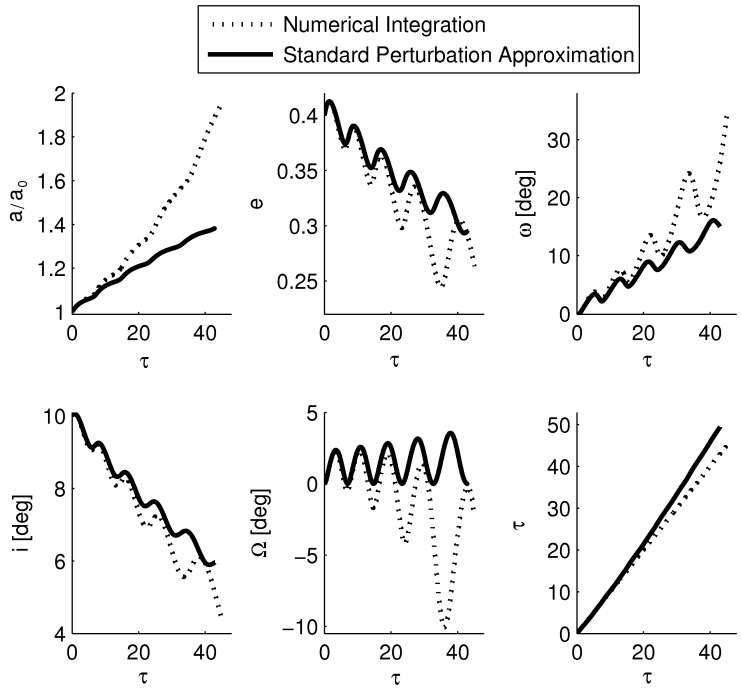

Figure 3.23: Parameters in time, $\varepsilon = 10^{-2}$

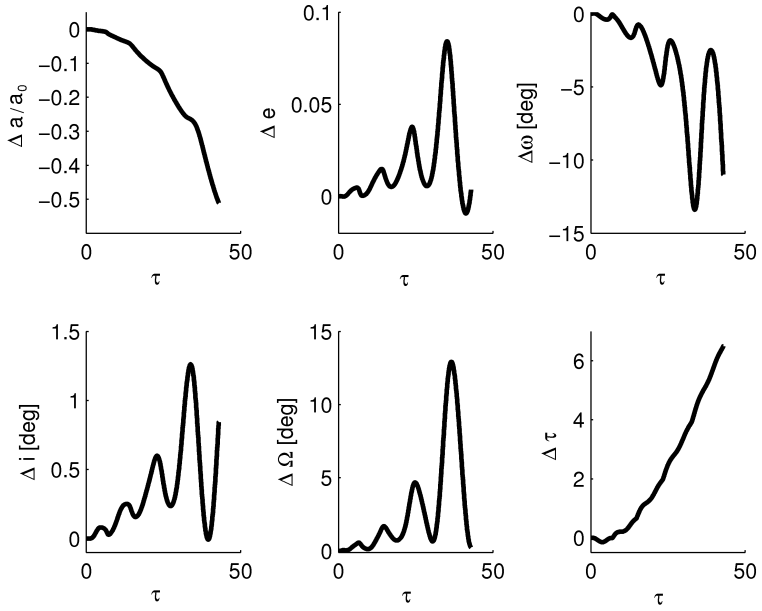

Figure 3.24: Error in time, $\varepsilon = 10^{-2}$

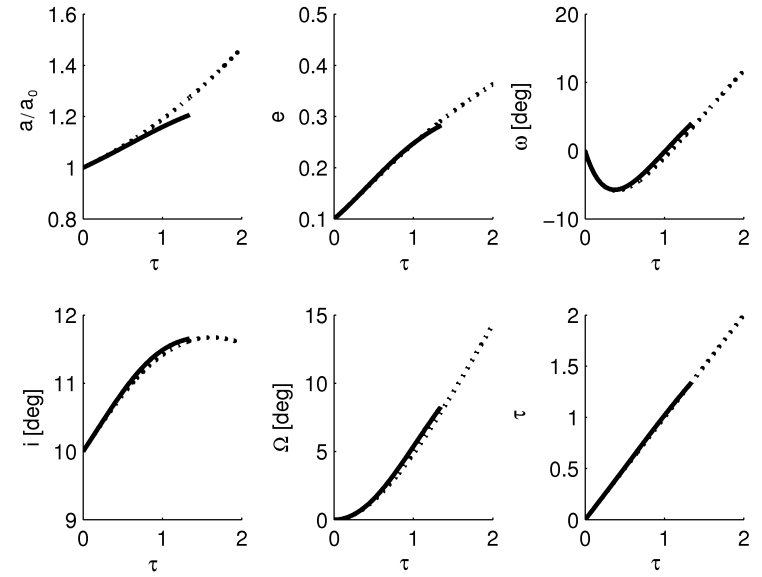

Figure 3.25: Parameters in time, $\varepsilon = 10^{-1}$

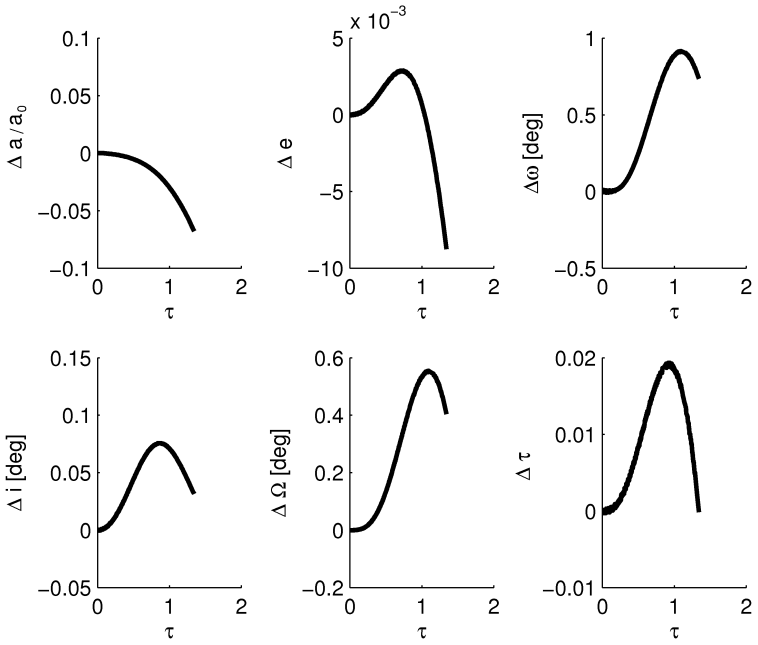

Figure 3.26: Error in time, $\varepsilon = 10^{-1}$

In figures 3.27 and 3.29 the trajectory is rappresented. It has the typical spiral shape, in case of $\varepsilon = 10^{-3}$ the approximation is satisfactory even for a large number of orbits, whereas for $\varepsilon = 10^{-2}$ the total angular travel must become lower in order to obtain a sufficient accuracy. In figures 3.28 and 3.30 the position error is rappresented with respect to time. Its value become high also for small values of time, this is caused, as said before for the two scale method, by the change in indipendent variable performed for the approach derivation. The time become a dipendent parameter and the error in its evaluation cause a rapid growth of the position error.

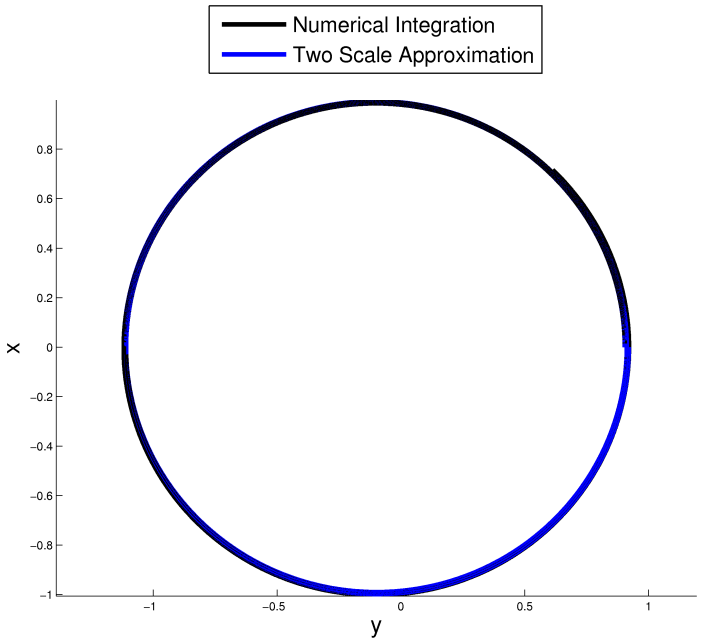

Figure 3.27: Trajectory, $\varepsilon = 10^{-3}$

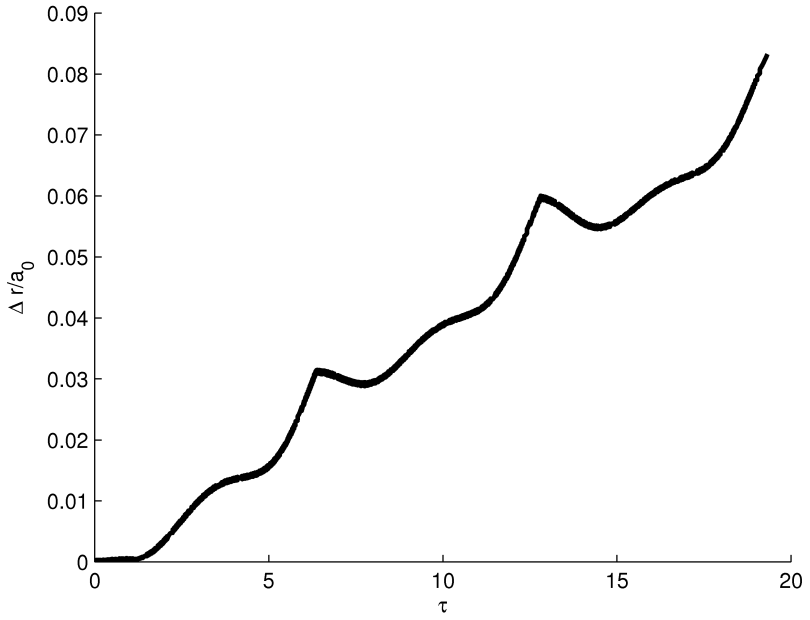

Figure 3.28: Error on the radius function of time, $\varepsilon = 10^{-3}$

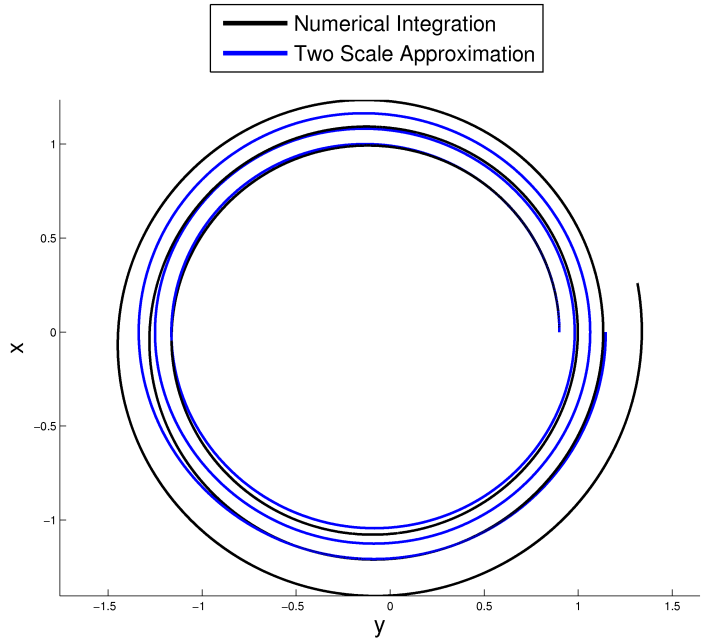

Figure 3.29: Trajectory, $\varepsilon = 10^{-2}$

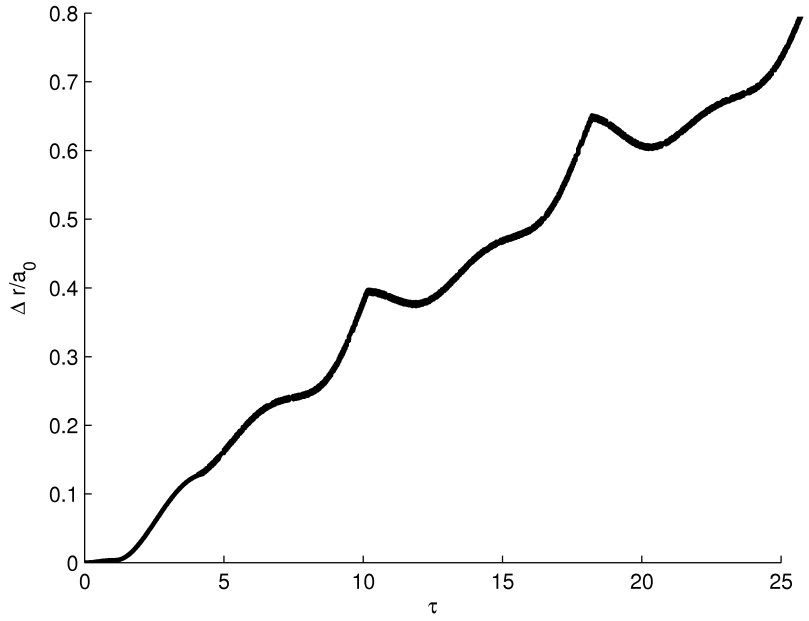

Figure 3.30: Error on the radius function of time, $\varepsilon = 10^{-2}$

3.3.2 Circular with Zero Inclination Case

For this case the angles for thrust direction have been chosen to be $\alpha = 20$ deg and $\beta = 20$ deg. As far as the evolution of orbit parameters and prediction error are concerned, the results are similar to those reported in the subsection for a generic elliptical inclined orbit. The approximation provides again satisfactory results in case of small acceleration, while its accuracy falls for large angular travels if ε grows. It has to be noticed that the formulation works well (even better due to the semplification seen in Chapter 2) also in presence of circular orbits with zero inclination thanks to the equinotial parameters formulation, which is not affected by singularities.

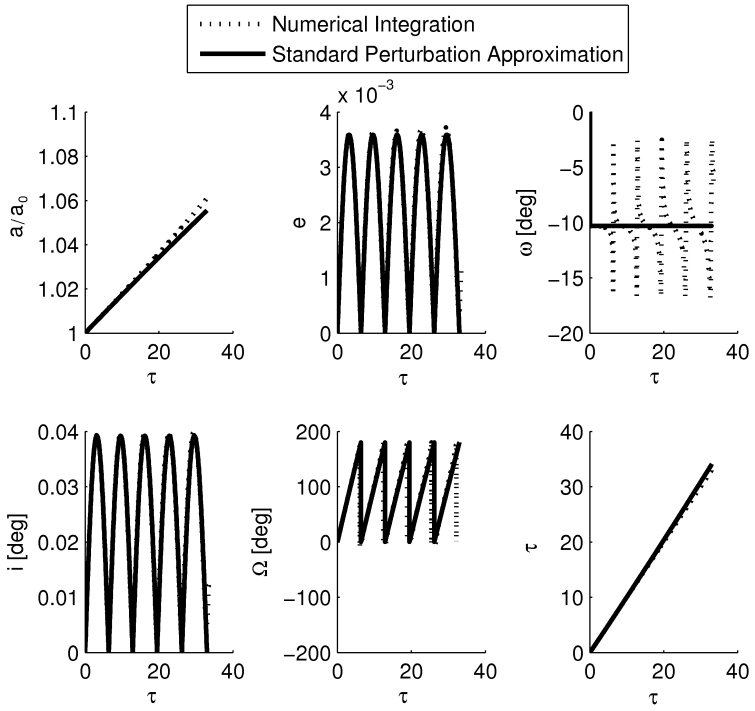

Figure 3.31: Parameters in time, $\varepsilon = 10^{-3}$

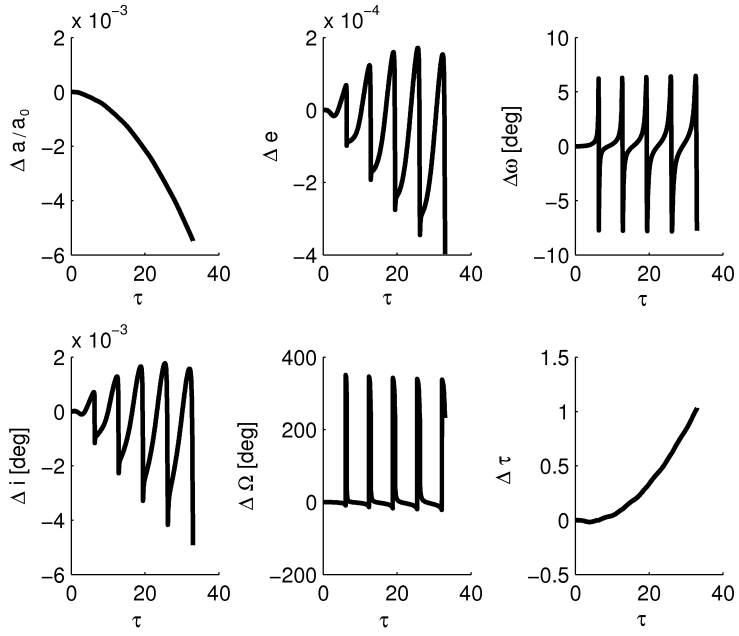

Figure 3.32: Error in time, $\varepsilon = 10^{-3}$

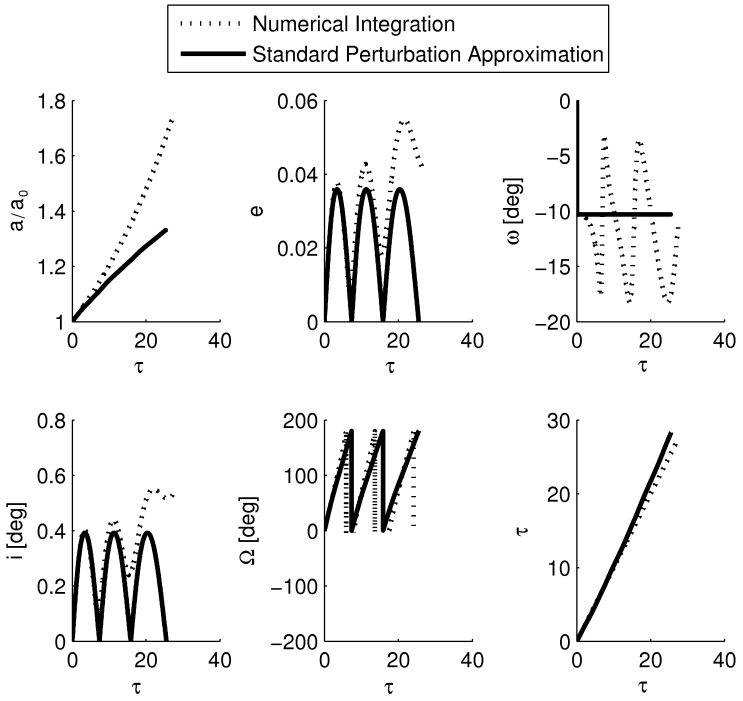

Figure 3.33: Parameters in time, $\varepsilon = 10^{-2}$

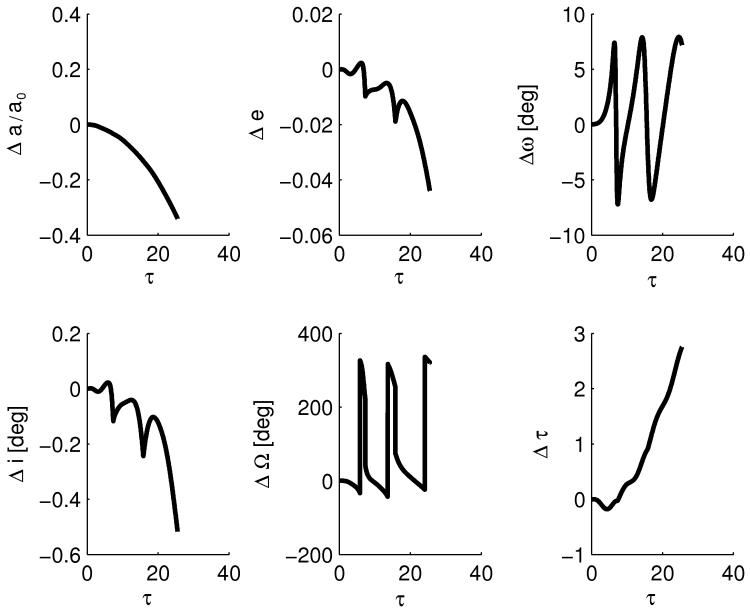

Figure 3.34: Error in time, $\varepsilon = 10^{-2}$

3.4 Accuracy Analysis

3.4.1 Error Analysis

In the previous two sections it was shown that the approximation error steadily grows with the amplitude of the trajectory arc, and that this growth becomes faster for higher values of the perturbation parameter ε, that is, the ratio of the acceleration produced by thrust and the gravitational one.

A detailed study on the evolution of the error as a function of the perturbation parameter, for various amplitudes of the trajectory arc thus rapresents an important step into the definition of limits within which the perturbation approximation provides an acceptable prediction of the variation of orbit parameters. As usual the reference solution for evaluating the error is obtained by means of a numerical propagation scheme based on Encke's method.

Figures 3.35, 3.36 and 3.37 show the trend of the average error (it's value is the mean one calculated considering the last approximated orbit) as a function ε. The values of ε vary between 10^{-3} and 1 for an angular trip less than one orbit (Fig. 3.35), while for one revolution ε ranges between 10^{-4} and 10^{-1} (Fig. 3.36). A smaller maximum value of $\varepsilon = 10^{-2}$ is considered for propagating the low–thrust trajectory for as many as 5 revolutions (Fig. 3.37).

It is clear how the error grows with ε and substantially there is an almost linear trend in logarithmic scale. The error can be mitigated if the angular trip is lowered.

The last figure, Fig. 3.37, clearly shows that, if a multiple revolution expansion is sought, the error on all the orbit parameters become large, even for relatively small values of ε. This means that the forward expansion must be truncated and the integrals updated, as it will be shown in the next section, or only very small value of the force delivered by the propulsion system can be considered. As far as practical applications are concerned, these very low values can be considered when dealing with solar sails [19].

Clearly when higher values of ε need to be considered, the angular travel must be considerably shortened (Fig. 3.35). For the same 1% accuracy on a, arclength of 90 deg at most can be considered for $\varepsilon = 0.1$, while 20 deg or even less can be propagated if ε grows close to unity.

This aspect, even if it is quite obvious, is very important since, as it will be seen later on, depending on the vehicle characteristics (mass and thrust) and the mission setting (distance from the main body), the value of ε could grow as high as 0.5, in these cases the small arc considered for assuming the approximation as valid must become smaller.

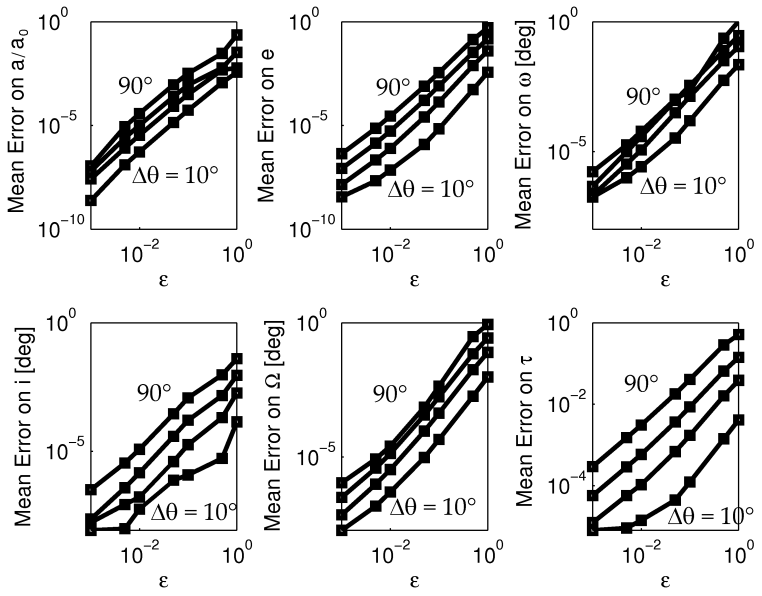

Figure 3.35: Mean error as a function of ε for different arc lengths $\Delta\theta = 10, 20, 45, 90$ deg.

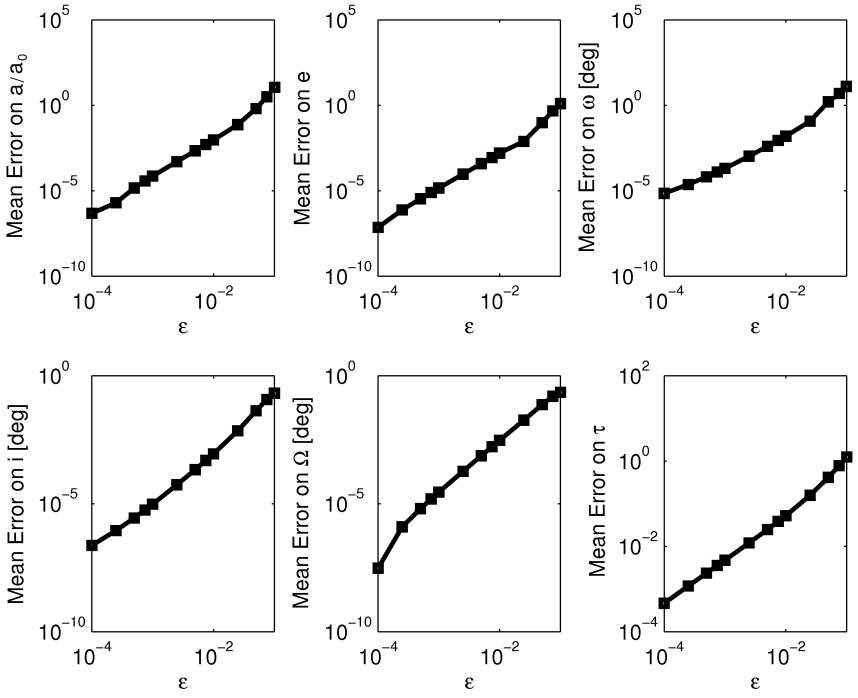

Figure 3.36: Mean error after one revolution as a function of ε

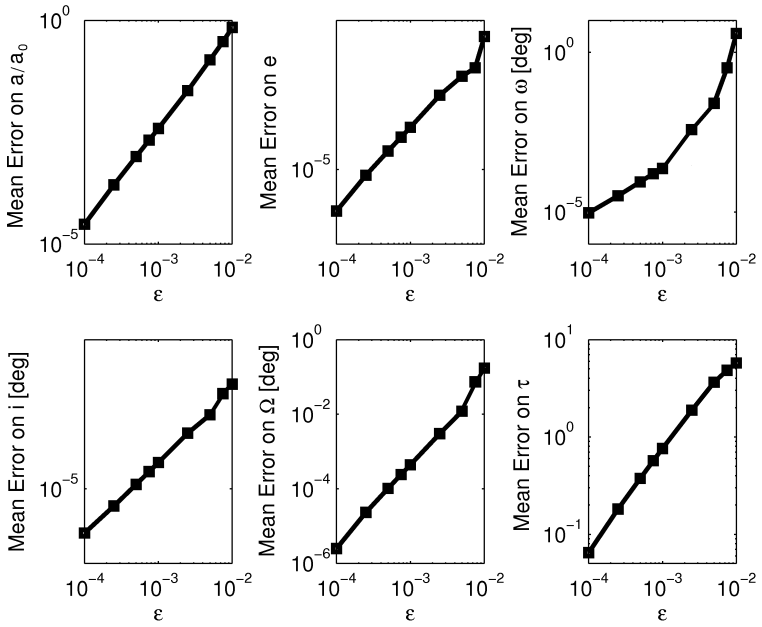

Figure 3.37: Mean error after five revolutions as a function of ε

3.4.2 Heliocentric ε variation analysis

From the previous discussion,the relevance of the values assumed by the perturbation parameter ε clearly emerged. As a consequence, its variation is here analyzed in an heliocentric scenario. The values of thrust and mass are taken from the datasheet of a modern mission, scheduled for 2014, the BepiColombo aimed at sending a probe towards Mercury. Different vehicles are still being considered for the mission, but the largest one has a mass of approximately 1200 kg and it is driven by electrical thrusters that deliver a force between 0.17 N and 0.34 N.

In Fig. 3.38 the trend for the gravitational acceleration μ/r^2 and acceleration ratio ε is reported as a function of the distance from the Sun. The curves are reported in Fig. 3.38 that are parametrized by thrust values equal to the maximum and minimum values considered. Figures 3.39 and 3.40 show isocontour lines for ε function of the distance. The two plots represent the limite cases, for maximum and the minimum thrust.

Quite obviously, when thrust is halved, the resulting value of ε is also reduced by 50% for the isocontour lines at the same distance from the Sun.

From all these observations it is clear how, for the considered vehicle, the value of ε would assume values up to some tenths in case a trip to Mars distance (approximately 1.5 AU) is dealt with. In this case the maximum angular trip allowed maintaining an erro betweeen 1 and 2% in the estimate of orbit parameters is between ten or twenty.

If trip to Jupiter distance is considered (5 AU) the value of ε grows to values which harm the validity of the considered expansion, which thus is no longer a valid orbit propagation tool, if a constant thrust is assumed.

In Figure 3.41 there is another kind of analysis, where for a limit value of $\varepsilon = 0.25$, and a mass equal to 1200 kg, the maximum distance as a function of thrust is determinated, which results in a value of ε confined below the limit set above. So it is clear how, for travel up to Jupiter distance, ϵ remains below $\varepsilon = 0.25$ if thrust is less 0.075 N. This values is very small, for ion thruster, if a constant thrust is available, indipendently of the distance from the sun, as if a nuclear reactor was used for producing the required electrical power. On the converse, the same value may be reasonable for solar sails, the thrust being inversely proportional to the distance form the sun, thus resulting in an almost constant value of ϵ.

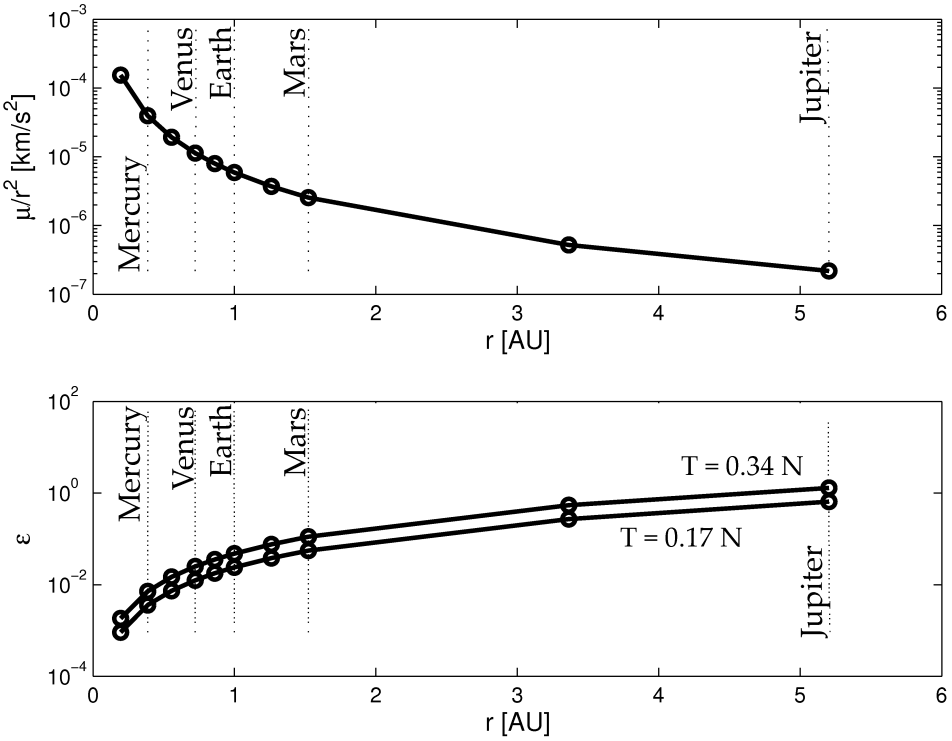

Figure 3.38: Gravitational acceleration and ε variation with distance, thrust varying $\in [0.17, 0.34]$ N and fixed $m = 1200$ kg

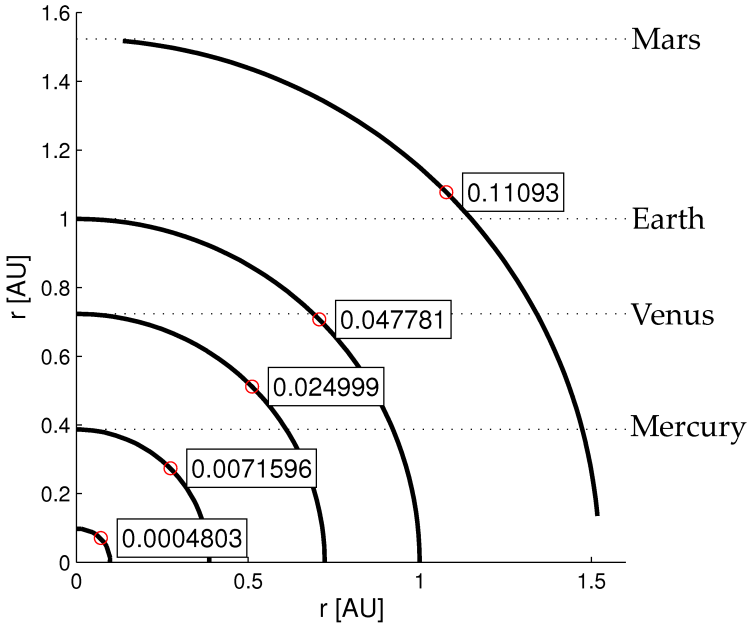

Figure 3.39: Iso-ε contour lines in 2D space, thrust 0.34 N and fixed $m = 1200\ kg$

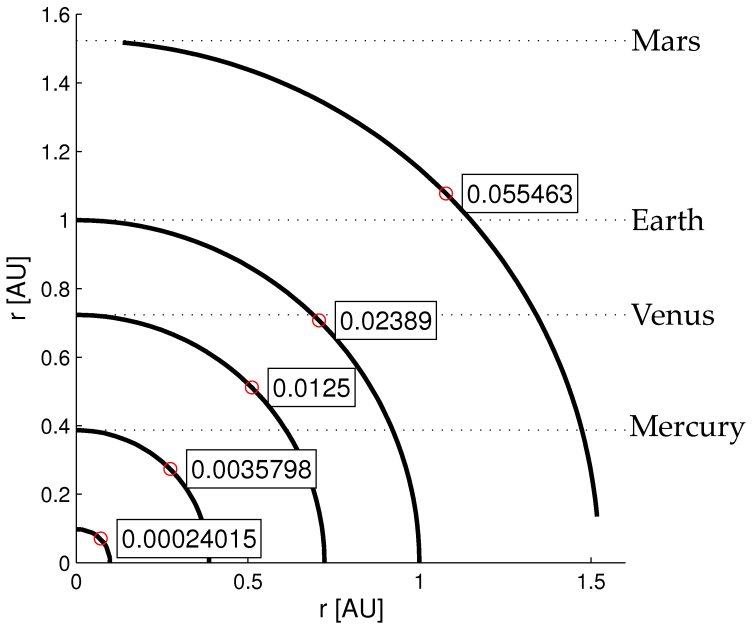

Figure 3.40: Iso-ε contour lines in 2D space, thrust 0.17 N and fixed $m = 1200\ kg$

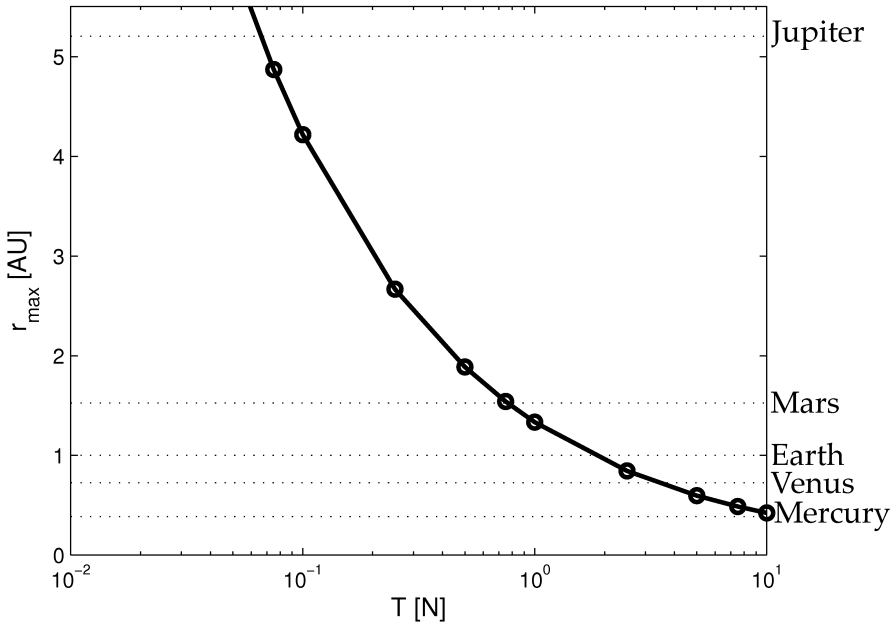

Figure 3.41: Maximum reacheble distance for a given $\varepsilon = 0.25$ and fixed $m = 1200\ kg$ function of thrust

3.4.3 Geocentric ε variation analysis

An analysis similar to that considered in the previous sub–section for a heliocentric scenario, is here reported in the framework of space missions (or mission segments) taking place within the sphere of influence of the Earth, up to Moon distance. In this case the same mass of the heliocentric analysis is assumed for the analysis, $m = 1200$ kg, with a single value for thrust, which is an intermediate value between the maximum and the minimum for the BepiColombo probe, i.e. 0.255 N. The limit on ε is the same as before. The results are qualitatively similar to those reported for the heliocentric case. Some characteristic distances like GEO and LEO orbit radii and the distance of the Laplacian Sphere of Influence are reported on the plots. It is clear how in this case also for high values of thrust the ε which result is quite low, so in case of geocentric trajectories the approximation can be extended to larger arcs, up to values as high as one/two orbits, while maintaining the error below some percentual points.

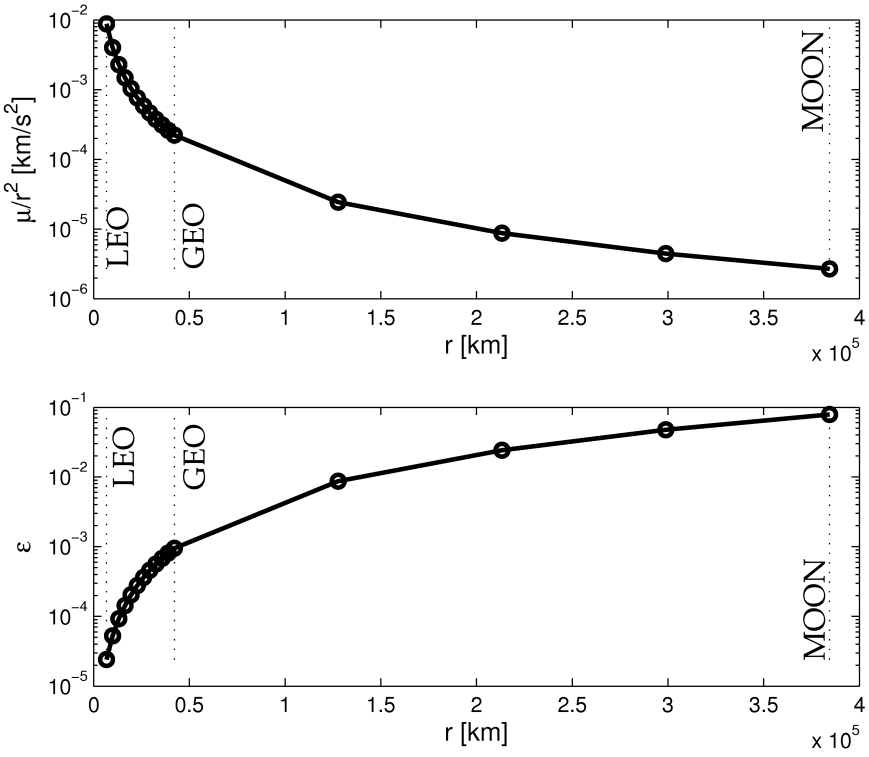

Figure 3.42: Gravitational potential and ε variation with distance, thrust 0.255 N and fixed $m = 1200$ kg

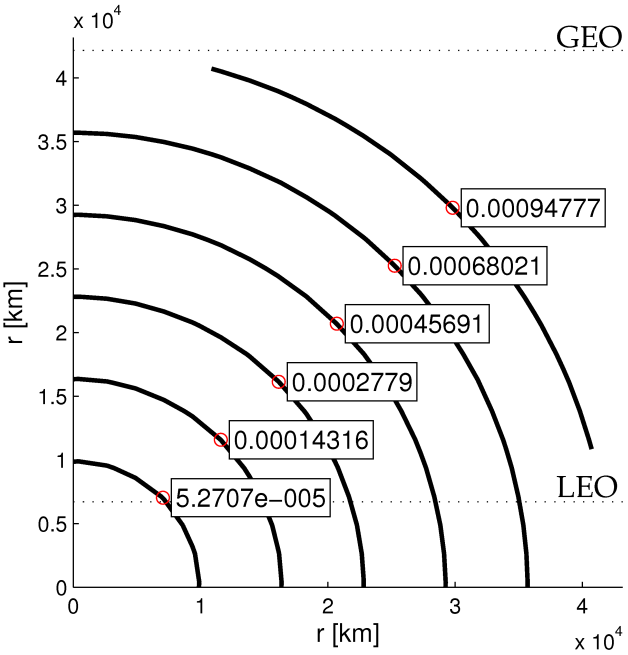

Figure 3.43: Iso-ε contour lines in 2D space, thrust 0.255 N and fixed $m = 1200\ kg$

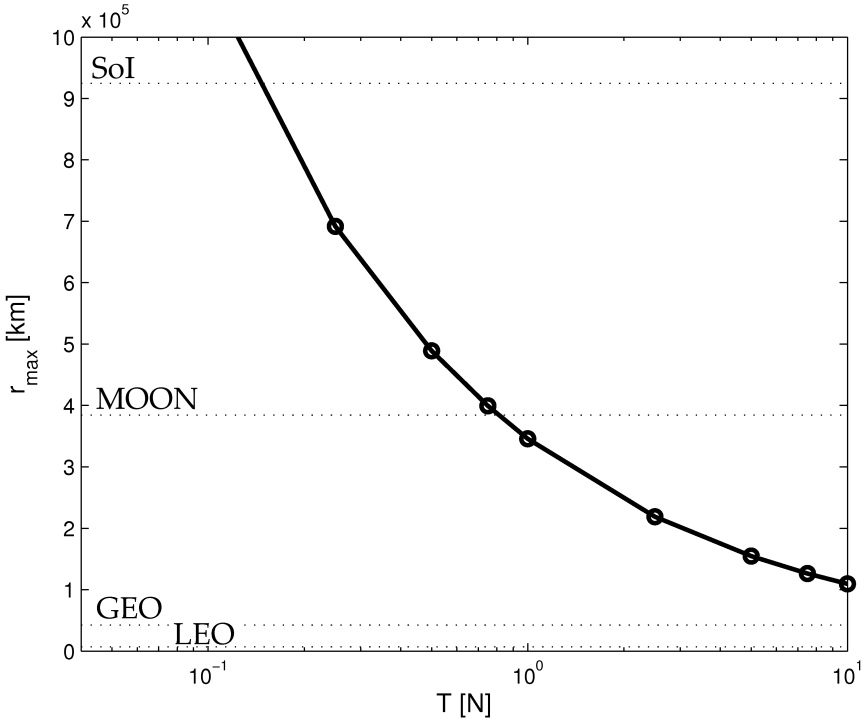

Figure 3.44: Maximum reacheble distance for a given $\varepsilon = 0.25$ and fixed $m = 1200\ kg$ function of thrust

3.5 Reference Comparison

One of the most recent work in the literature which performs a similar search for analytical relations for describing the motion under low–thrust propulsion is the paper by Scheeres and Hudson [18]. They use a Fourier expansion and by means of this tool they approximate the thrust profile first and then introduce it in Gauss's variational equations written in the classical parameter formulation. By analitical integrating they obtain the rate of change of each parameter and these six quantities are function of a restricted number of the Fourier expansion coefficients used for approximate the thrust profile.

Beside the high number of relevant interesting features of the work, like the speed of the algorithm and the semplicity with which it can approximate every kind of thrust profile with variable magnitude and direction (this aspect steams directly from the Fourier series capability), there are two major limits:

- It is capable to capture only secular (long term) variations, without taking into account short term ones;
- It can not deal with cases with zero eccentricity and inclination, close to zero;
- Only periodic variation of thrust components can be dealth with.

In spite of this limits for its application, this approach represent and interesting reference for the performance of the perturbative approach. A comparison between the two solutions to the propagation of orbit parameters under low–thrust propulsion will thus be considered in the sequel. This will be done on the basis of the examples reported in [18] where the thrust profile is assumed to be defined as follows

$$T(E) = \begin{cases} 10^{-7} & \text{if} \quad E \in \left[\frac{\pi}{2}, \pi\right] \quad \wedge \quad E \in \left[\frac{3\pi}{2}, 2\pi\right] \\ 0 & \text{if} \quad E \in \left[0, \frac{\pi}{2}\right] \quad \wedge \quad E \in \left[\pi, \frac{3\pi}{2}\right] \end{cases}$$

where E is the eccentric anomaly and the direction of the thrust is only trasversal.

For such kind of profile a comparison between four different methods will be considered

1. Numerical Encke's Method;
2. Scheeres's Method;
3. Two–Scale Perturbative Approach;

4. Standard Perturbation Method.

The trend of the three relevant parameters for the planar example considered (semimajor axis a, eccentricity e and argument of the perigee ω) is rappresented in the figure 3.45. The expansion is performed for ten orbits.

As clearly shown, the Fourier–based approach catches well the long period variation, exactly like two scale method (but the latter one demonstrate less accuracy for the argument of periapsis). The standard perturbation is able to catch also the short period term and it remains really close to the numerical integration result. This fact can be observed clearly in the second picture where the difference between them remains confined below 10^{-2}-10^{-3} even for ten orbits .

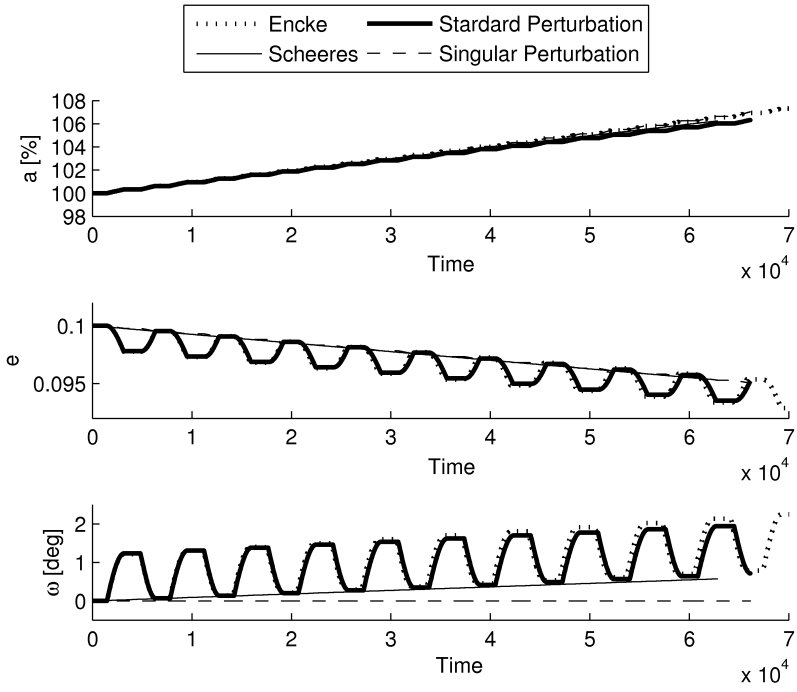

Figure 3.45: Parameters in time

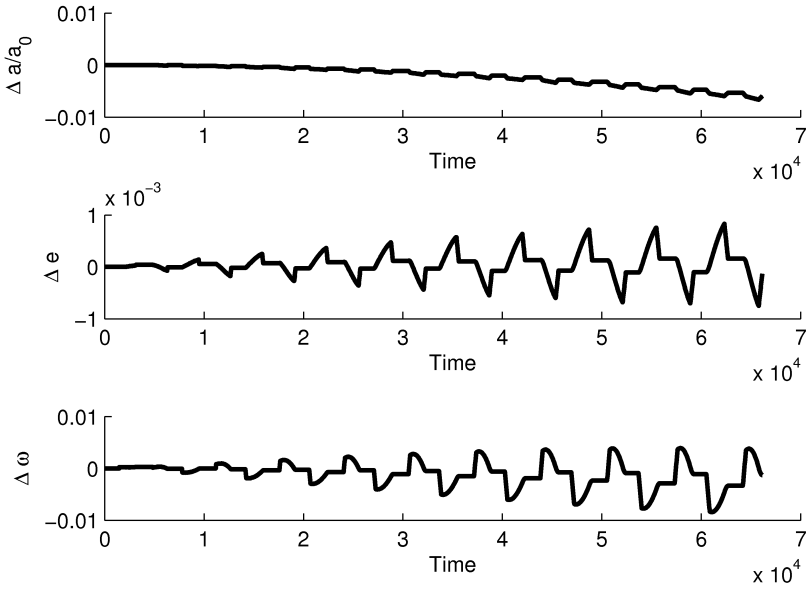

Figure 3.46: Error in time, standard perturbation

3.6 Computational Time Evaluation

In this section the computational time of the expansions derived in this work is analyzed. As a reference point the numerical integration performed by means of the Enkce method is used. The time posted in the following tables is evaluated by means of the `tic toc` Matlab©routine.

In Table 3.2 the comparison between two–scale method and an accurate numerical integration scheme is presented.

ε	$t_{num}\ [s]$	$t_{twoscale}\ [s]$	$t_{twoscale}/t_{num}$
10^{-3}	15.036634	0.256453	0.01705
10^{-4}	14.588808	0.238075	0.01632

Table 3.2: Computational time comparison for two–scale method.

In case of $\varepsilon = 10^{-3}$ the value of computational time for the analytical solution is 1.705% of the numerical one while for $\varepsilon = 10^{-4}$ the percentage is 1.632%. This ratio can be considered not related to the number of orbits considered and from the value of the adimensionalized acceleration ε.

In Table 3.3 the comparison between standard perturbation method and an accurate numerical integration scheme is presented.

ε	$t_{num}\ [s]$	$t_{perturbation}\ [s]$	$t_{perturbation}/t_{num}$
10^{-3}	53.774197	5.060510	0.0941
10^{-4}	52.776370	5.995314	0.1136

Table 3.3: Computational time comparison for standard perturbation method.

In case of $\varepsilon = 10^{-4}$ the value of the computational time used for obtaining the analytical solution is 11.36% of the numerical one while for $\varepsilon = 10^{-3}$ the percentage is 9.41%. Again, this ratio can be considered not related the number of orbit simulation and from the value of the adimensionalized acceleration.

It must be underlined that the comparison here presented is truthful if one wants to obtain from the analitical approximation the same result given by the numerical one, that is the orbit propagation with the calculation of the intermediate points between the initial and the final one. But whereas the numerical one must always do this task for obtaining the final position and velocity, the analytical approximation could avoid the evaluation of all

the intermediate points and calculate directly the final required values. This is true is is not need to update the parameters values during the whole propagation, if it is needed, some internal points are needed.

In Table 3.4: are presented the result for the comparison of the analytical approximations derived in this thesis with the analytical approximation by Scheeres [18]. Scheeres's method is about two times faster than the twoscale one and many more times respect to the standard perturbation method, but with respect to this one it does not capture the short period terms, as underlined before. In addition, it must be note that a great amount of computational cost for Scheeres's method is not taken into account here, that is all the calculations needed to obtain the Fourier coefficients for a given thrust profile.

$t_{scheeres}$ $[s]$	$t_{twoscale}$ $[s]$	t_{pert} $[s]$	$t_{scheeres}/t_{twoscale}$	$t_{scheeres}/t_{pert}$
0.230017	13.184149	0.492038	0.4675	0.01744

Table 3.4: Computational time comparison for different approaches.

3.7 Perturbative Expansions Applied to Optimization Methods

In this section a preliminary application of the perturbative expansions to a Direct Optimization Method is outlined. The most important benefits that this technique can introduce. An overview on optimization methods is reported in Appendix C.

The most important aspects in which the introduction of the perturbative expansions present significant advantages are:

- Orbit propagation/trajectory computation from the given initial values forward to the final point of the sub–element by means of an analytic expansion instead of numerical orbit propagation, thus increasing computation efficiency;
- The possibility of using Automatic Differentiation, for analiticaly computing the gradients in the optimization process, improving its convergence rate.

Both these aspects are analyzed in a dedicated subsection; the result is a new formulation for a sub–element by means of which the trajectory can be discretized that should provide a great reduction of computational time, as shown in the next section where results are reported.

3.7.1 Orbit Propagation

Direct optimization methods are usually implemented by means of the direct transcription technique which implies the multiple shooting approach: the trajectory is discretized in finite sub–elements and the initial points of each subelement are guessed. Then, togheter with constraints enforcement (boundary conditions, state equations, etc.) the matching conditions at the nodes must be introduced (the final value of each subelement for a variable must be the same of the initial value on the following one).

Usually the indipendent variable is time but since the analytical relations found in the previous sections for the trajectory description in case of low–thrust are parametrized by means of the mean longitude L this will be assumed to be the indipendendent variable. Time will assume the role of a common dependent variable, varying according to its own approximated analytical relation.

For discretizing the trajectory one can look at the total angular trip an divide it in a given number N of sub arcs. It is clear how, once the initial conditions in each node are given (state and control variables $\boldsymbol{x}$ and $\boldsymbol{u}$), a numerical propagation which can determine the trend for all the different variables in that sub–element is required. So a natural way for using the approximated relations that describe such a variation under the action of a low–thrust propulsion system is to use them for the orbit propagation (insted of the numerical integration of the dynamic equations), since only the initial conditions on state variables (five orbital parameters and time) and the control variables (magnitude of the acceleration and two angles) must be known.

At this point a key issues rises: the value of N, the number of subarcs, which choice is of critical importance.

First of all the total number of optimization variables is proportional to N; if one consider six state variables plus three control variables the number of variables y is given by

$$y = 9N$$

As an example, when an Earth-Mars trajectory is considered, with an angular trip of three and a half revolutions, considering sub–elements of 21 degrees each, one obatins

$$N = \frac{3.5 \cdot 360}{21} = 60$$

so that the number of variables is

$$y = 9 \times 60 = 540$$

which is a very large number of optimization variables;

At the same time its value is linked to the error allowed by the perturbative expansion. It has be seen in the previous sections that the error depends on the value of the ratio of the low–thrust acceleration to the gravitational one, and if this value grows the arc length must be reduced in order to limit the error of the approximation.

The orbit propagation can be done in different ways, for example it can be forward expanded or centered depending on how the initial state is expanded. In the first case (forward expansion in Longiture and/or time), given the set of initial values on one subelement, the expansion is performed only forward in L (see figure 3.47);

In the second one (centered expansion), given the set of initial values on one subelement, the expansion is performed backward and forward in L (see figure 3.48) for half of the arc amplitude.

With the centered expansion propagation is done for one half of the total travel in both directions for the same total subarc angular amplitude, so that a better approximation, i.e. a smaller error, is expected.

Another modification can be the introduction of a mix of the two tecniques, namely two scale and standard perturbation, the former (less affected from the drift due to the epsilon value) valid for in plane orbital parameters and the latter for the out of plane ones plus time. Although interesting in principle this aspect will not be investigated in this work.

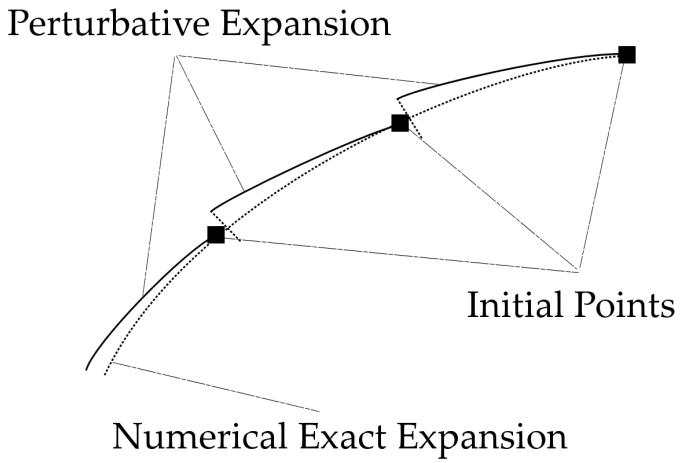

Figure 3.47: Forward Expanded

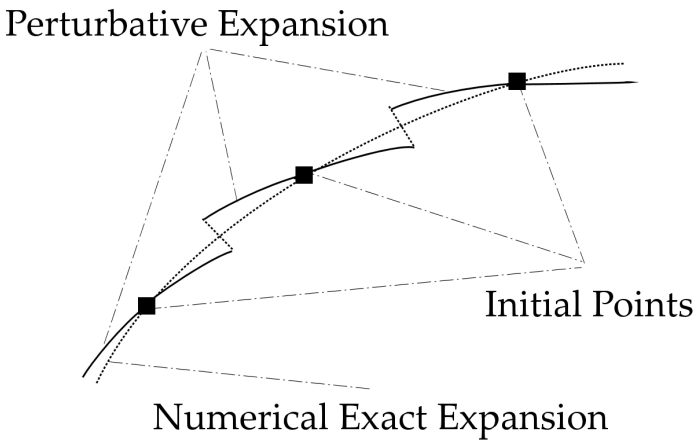

Figure 3.48: Centered

3.7.2 Automatic Differentiation

In order to find the optimal solution, derivatives of the performance index and constraints with respect to the parameters need to be evaluated. This task is usually performed numerically and obviously this is a computationally demanding process and, morover, it introduces numerical inaccuracies in the evaluation of the gradients, which reduce convergence rate.

A different approach to be used when analytical expression for functional dependencies are available, relies on automatic differentiation procedure which consist in obtaining analytically the derivatives, thus eliminating a numerical process and related inaccuracies. There are a lot of numerical algorithms that can performautomatic differentiation, but their speed depends on the particular form of the functions and each of them must be properly interfaced with the rest of the algorithm. It should be noted that the computational cost for automatic differentiation is paid only in the first iteration, once all the derivatives are evaluated their value remains fixed in the following interations.

3.8 Optimization Methods Results

In this section the main features and results of a preliminary application of the standard perturbative expansions to a direct optimization method is presented. The solution method is based on the constrained otpimization routine `fmincon` available in Matlab©The main parameters for the implementation

are:

- N: number of the finite arcs in which the trajectory is discretized;
- J: performance index, which is identified in the total ΔV for the considered mission;
- $\boldsymbol{c}$: a vector of constrains.

There are four kind of constraints for the problem:

- initial and final position and velocity (boundary conditions);
- matching constrains at the nodes;
- time constrain;
- acceleration contraints (related to the thrust limit).

In order to test the optimization algorithm an Earth-Mars trasfer like was simulated. Transfer and vehicle data assumed for the problem are:

- Circular orbits radii: from 1 to 1.5 AU;
- Maximum thrust: 0.5 N;
- Spacecraft mass: 1000 kg;
- Maximum transfer time $1.25802 \cdot 10^3$ gg = 3 years 5 months 13 days 0 hours 35 minutes 46 seconds, (which was the result obtained from another optimization technique);
- Initial guess: biimpulsive transfer with two additional revolutions, for a total angular trip of 2.5 revolutions;
- Number of finite elements (subarcs): $N = 80$, resulting in an angular trip for each element of 11.25 deg;
- Orbital propagation implemented with the analytical expansions using the centered method, no automatic differentiation;
- Contrains tollerance 10^{-8}

For these considered data, an acceleration ratio of $\varepsilon = 0.084308$ is obtained;

The optimization results are presented in the Figures 3.49, 3.50 and 3.51. Fig. 3.49 shows the trajectory, Fig. 3.50 the variation of the equinocial orbital elements as a function of L and, finally, Fig. 3.51 the thrust profiles with, respectively, magnitude, in plane and out of plane angles for each subelement.

In the Table 3.5 is presented the value of the performance index which is the total ΔV compared with the one obtained from the hohmann transfer.

	$\Delta V_{hohmann}\ [km/s]$	$\Delta V_{lowthrust}\ [km/s]$
ΔV	5.41048	5.37851

Table 3.5: ΔV values for the mission.

For the error with respect to a complete numerical solution and the computational time evaluation[2] a simple case is taken as another example, it is a simple orbit raise in LEO setting, with tangential thrust of low magnitude, also for this case 2.5 revolutions are taken into account, both the values are reported in the table below, where 'E-M' means Earth-Mars transfer, 'num' means numerical result and 'exp' means expansion result: it can be seen that the introduction of the perturbative expansions allow to save a lot of computational time, with an acceptable error also for a setting (the Earth-Mars one) with a value of the parameter ε not so small. It has to be underlined that the use of the centered expansion allow, for the same number of elements, to reach two/three times the precision reacheble in case of forward expansion.

	LEO_{num} [s]	LEO_{exp} [s]	E-M_{num} [s]	E-M_{exp} [s]
Time	1.71	0.0013	0.6679	0.0424
Time ratio	1.00	$7.602 \cdot 10^{-4}$	1.00	0.0634
Error	–	$\mathcal{O}\left(10^{-5}\right)$	–	$3.867 \cdot 10^{-3}$

Table 3.6: Computational time comparison.

[2]The time is relative to the numerical propagation.

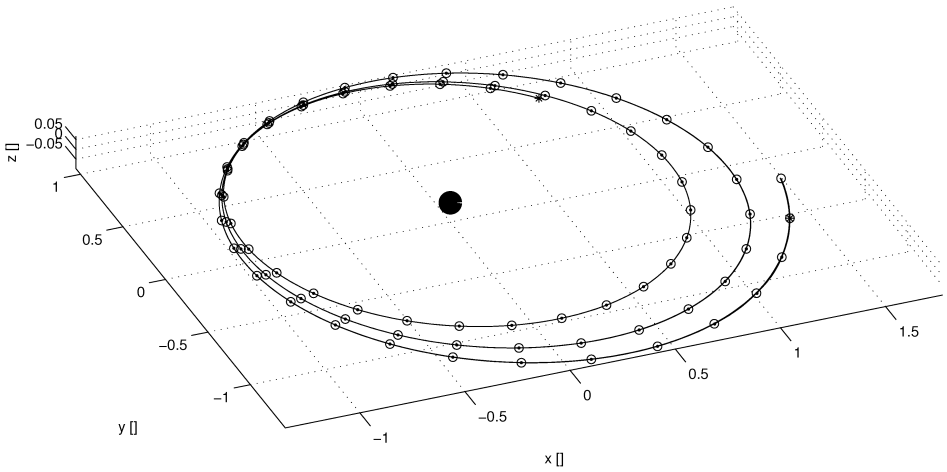

Figure 3.49: Optimized trajectory

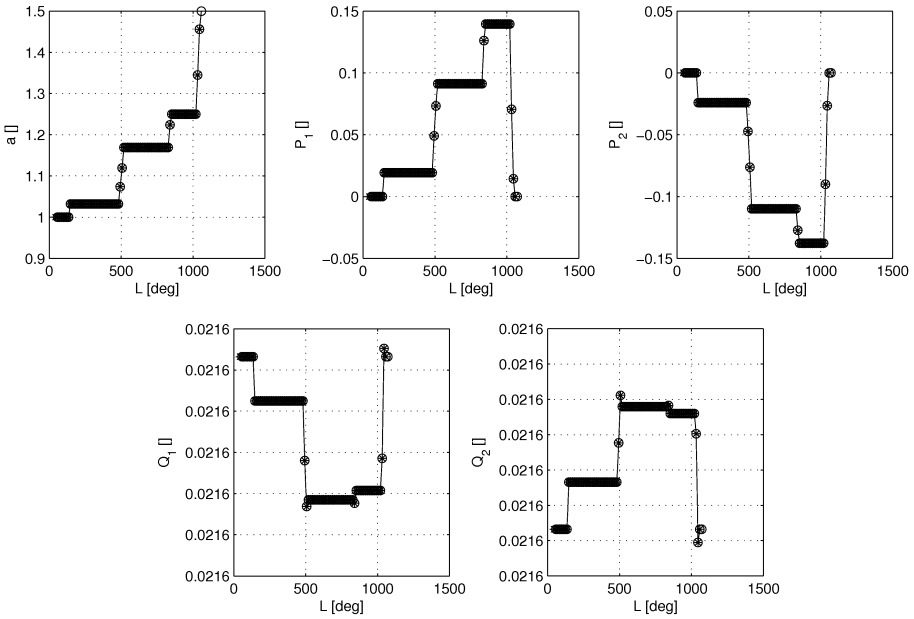

Figure 3.50: Parameters variation: a, P_1, P_2, Q_1, Q_2

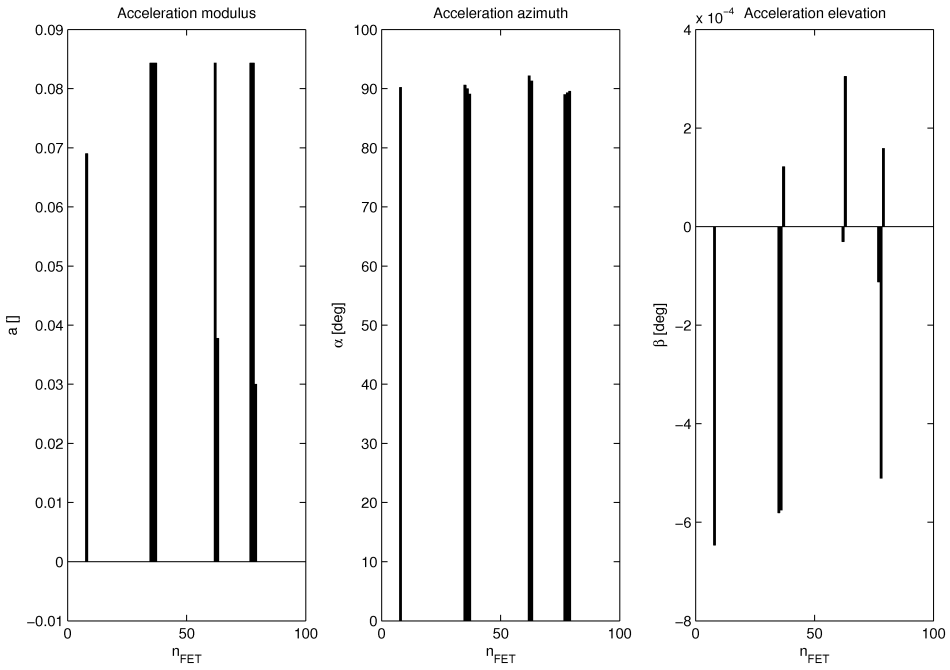

Figure 3.51: Thrust profile: magnitude, azimuth α and elevation β

Chapter 4

Conclusions

4.1 General features

Two different analytical approximations for low-thrust trajectory arcs were developed and analyzed, both based on a perturbative approach where an expansion in terms of angular coordinate variable is considered. The first approach is based on a multiple–scale perturbative expansion of the fundamental equation of Astrodynamics, written in polar coordinates, with respect to the anomaly, θ. The second method introduces a standard perturbative expansion in terms of mean longitude L of the Gauss's variational equations.

The two scale method is simpler, but it is limited to planar analysis, while the second one is more complex, but allow to consider fully threedimensional motion.

The results obtained by means of both expansion in evaluating the evolution of orbit parameters under low–thrust are quite satisfactory, especially when the standard perturbation method is considered, as it is capable of capturing both short and long period terms and it is not affected by siingularities induced on classical orbital parameters by circular and zero inclination orbits, as an equinocial elements formulation was derived. The physical consistency of the two formulations was proved by means of several simulations and a comparisons with results obtained by means of an accurate numerical method. Also another analytical formulation, based on a Fourier series expansion, was considered in the framework of the test-benchmarks, which proved to be more efficient in terms of computational time, but less general as it can accomodate only cases with periodic variations of thrust and it suffer singularities for circular and zero–inclination orbits. Nontheless the savings in computational time with respect to numerical solutions, are considerable, if one wants to propagate the orbit under the action of a low–thrust propul-

sion system, both analytical formulations allowing for a reduction of CPU time between 90 and 98%, when one wants to obtain all the points along the arc. But the method allows also for the direct estimate of the final values at the end of a given arc, in which case the analytical approximation is simply not comparable with a numerical solution which always requires significantly smaller integration steps.

There are some limitations that restrict the applicability of both the formulations. In general, the ratio ε between spacecraft and gravitational acceleration must be limited, in order to maintain the error on the propagation of orbit parameters into acceptable limits. For larger values of ε the arc–length of the expansion must be significantly reduced. As for the two–scale approach, more specific limitaiton are represented by:

- thrust is assumed constant in magnitude and its direction either aligned with velocity or fixed with respect to the radial-transversal reference system;

- the derivation was performed for the planar case only.

This latter aspect is particularly relevant, inasmuch as if considerably restricts the type of orbit transfers that can be described by this approach, ruling out every mission or mission segment that requires an orbit plane change.

Also for the standard perturbation: thrust magnitude is assumed constant, with a thrust direction fixed with respect to the radial–transversal–normal reference system, but as for the rest, the approach is absolutely general.

It should be noted that these latter limitations (namely the assumption of constant magnitude and direction of thrust) do not seriously harm the relevance of the proposed approach, as long as it is possible to discretize the trajectory and assume piecewise constant control for each sub–arc, the amplitude of which is determined in such a way that the resulting error remains bounded within acceptable limits.

A detailed error analysis has been performed and it has be seen that precision is stricly a function of the acceleration ratio ε (thrust over gravitational) and of the approximated arc length. The value of ε was investigated in two tipical contexts, geocentric and heliocentric, for a modern space vehicle configuration. With current tecnology, the value of ε in geocentric scenarios remains quite low, so relatively large arcs (of several tens of degrees) could be used also for a distance equal to the Laplacian sphere of influence. In case of remote mission (as far as Jupiter) the arc approximation must be reduced due to the high value of the acceleration ratio and it should be ten or twenty

degrees long at most. Of course in case of larger distance, the more traditional impulsive approximation could be used when dealing with electrical propulsion, since the time of the burn becomes negligible with respect to the orbital period. In this case an analytical approximation for large low–thrust trajectory arcs is not needed anymore.

4.2 Preliminary Application to a Direct Optimization Problem

The analitical relations found by means of the standard perturbation approach were applied to a direct optimization problem, where their task was essentially that of replacing the numerical integration of the equations of motion in the trajectory propagation step. Another possible use is the automatic differentiation for the evaluation of performance index and constrains gradients with respect to optimization variables, which was not investigated in the present study. Two methods for use of the analytical relations instead of the numerical integration where considered: a forward expansion and centered one. In the first one the expansion is performed only forward with respect to the angular variable, L, while in the second one the control point was set at half of the arc–length, so that backward and forward expansions were necessary for determining the values at the edges of each discretization arc. This second technique allows for a significantly better accuracy in the prediction, up to two or three times a higher precision with respect to the forward expansion formulation, for the same total arc length.

These techniques allow for a sizable reduction of computational cost of the whole optimization process with respect to the use of a numerical integration algorithm. A straigthforward application of perturbative expansions is thus the development of a fast algorithm which can help in selecting good approximated optimal solutions which can be used as first guesses for a fully numerical more accurate direct optimization algorithm.

Future Work and Extension of the Approach

Further steps for the research on the application of perturbative expansions to low–thrust orbit propagation can be outlined as follows:

- **Implementation of an effective automatic differentiation**: the availability of an analytical expansion for orbit propagation allows one to formulate analytical expression for performance index gradient and constraint Jacobian matrix, with the aim of a fruther signigicant reduction of the total computational time for the optimization procedure.

- **Implementation of a mixed tecnique**: two–scale and standard perturbation techniques can be used in a mixed approach which could allow the same precision with the use of larger sub–arcs, due to the lower drift caused by large ε values in the two scale approximation, which propagates secular terms only.

- It is also possible to figure out a fully three–dimensional formulation for the two–scale approximation, in order to broaden its applicability to a larger class of problems, or to derive a second–order formulation for the standard perturbative approach, in order to improve its accuracy and allow a discretization of low–thrust trajectory based on larger sub–arcs. Both these tasks appear to be extremely demanding from the mathematical stand–point, but they would provide a truly significant improvement to the potential applications of perturbative expansions in the framework of space missions analysis and design.

Appendix A

Fundamentals of Space Flight Dynamics - Orbital Mechanics

In this chapter will be resumed all the most important astrodynamics concepts of the two body orbital mechanics which is the most common setting for studying the motion of a spacecraft in space.

A.1 Two Body Problem

In the preliminary study of an orbital transfer one can assume, in order to simplify the problem, the presence of interaction only between two bodies: the main (primary) body and the space vehicle itself, assuming the other influences (those caused by the other celestial bodies, the solar radiation, the asphericity of planets, etc.) negliable.

Kepler Laws

The first step towards the modern Astrodynamics it is used to be identified with the studies of Johannes Keplero, who between 1609 and 1619 wrote down the famous Three Laws, derived from the analysis of data on the planet movement, previously collecter by the studious Tycho Brahe. The three Kepler Laws can be resumed as follows:

First Law:*Each planet orbit is an ellipse with the Sun positioned on its focus (cf. fig. A.1);*

Second Law: *The position vector from the center of the Sun to the center of the Planet sweeps out equal areas in equal times (cf. fig. A.2);*

Third Law: *The sqare of a planet period is proportional to the cube of its mean distance from the Sun.*

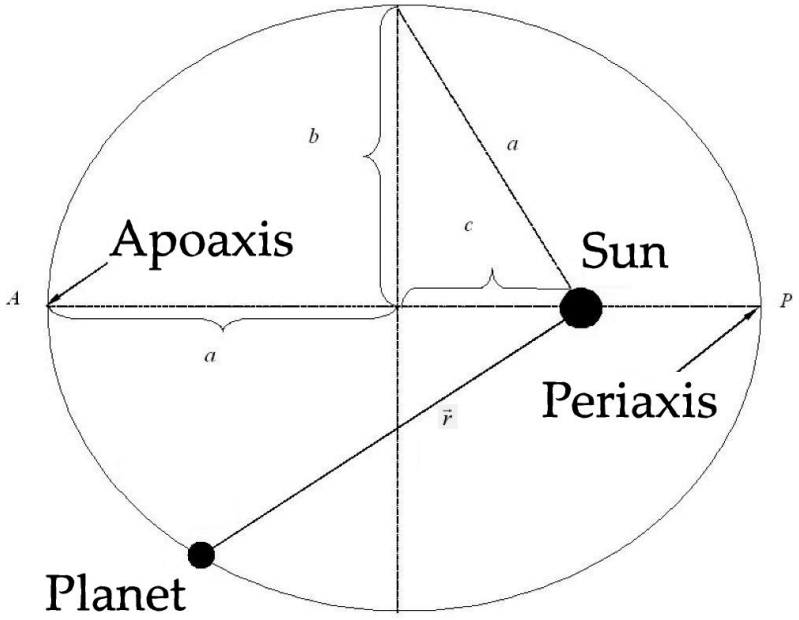

Figure A.1: First Kepler Law

The First and the Second Law can be generalised to those bodies that moves with a velocity greater than the so called escape velocity, like some comets do, considering open orbits, parabolic and hyperbolic ones, with the Sun on the focus of the curve. This means that every possible trajectory in the two body problem is a conic section. The Second Law express the conservation of the angular moment: the distance between the planet and the Sun will vary with the angular position of the planet and with a proportionality law the velocity will vary too. The same kind of generalization cannot be applied on the third law, infact the open orbits are not periodic.

Newtonian Law of Motion

Some decades after Kepler, Isaac Newton introduced (year 1687) the Three Laws which nowadays are the classical mechanics foundations:

First Law: *Every body continues in its state of rest or of uniform motion in a straight line unless it is compelled to change the state by force impressed on it.* (This Law is valid once one can identify an inertial reference system.)*;*

Second Law: *The rate of change of momentum is proportional to the force impressed and is in the same direction as that force;*

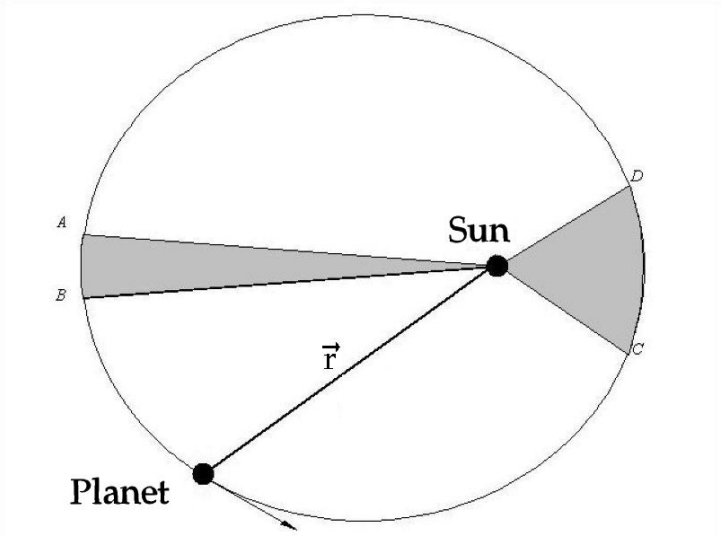

Figure A.2: Second Kepler Law

Third Law: *To every action there is always opposed and equal reaction.*

The Second Law can be expressed, assuming a constant mass, with the following relation:

$$\vec{\mathbf{F}} = m\vec{\mathbf{a}}$$

with $\vec{\mathbf{F}}$ the resultant of applied forces and $\vec{\mathbf{a}} = d\vec{\mathbf{v}}/dt$ absolute acceleration measured in an inertial reference system.

Newton's Law of Universal Gravitation

Besides the Three Laws of Motion, Newton introduced the Law of Universal Gravitation, this law in the Two Bodies Problem of mass respectively M e m, gives the magnitude and the direction of the force of attraction between the two bodies; this force is directly proportional to the product of the two masses and inversely proportional to the sqare of the distance between them:

$$\vec{\mathbf{F}} = -G\frac{Mm}{r^2}\hat{\mathbf{u}}$$

where $G = 6.67310^{-11} m^3/Kg\ s^2$ Universal Gravitation Constant and $\hat{\mathbf{u}}$ versor directed as the line joining the center of the two bodies, poining to the smaller one. This law is exaclty valid not only for bodies considered as a point, but also for those with a spherical symmetry and even for any form of planet if considered at a distance so big that the body dimension can be considered negliable respect to the distance.

Equation of motion

If n-bodies are considered, with the only gravitational force acting between them, writing down the law of motion for each of them one derives the so called n-bodies problem. It is modelled in matematical term by a system of n non linear second–order vectorial differential equations, it can be written down using the Law of Universal Gravitation (in the real case beisdes these terms there are others terms linked to the different kind of perturbations) and the second law of motion, with the position rappresented by the vector $\vec{\mathbf{R}}_i$. If n became grater that 2 it is impossible to obtain analytical solution of the problem and it is necessary to procede numerically.

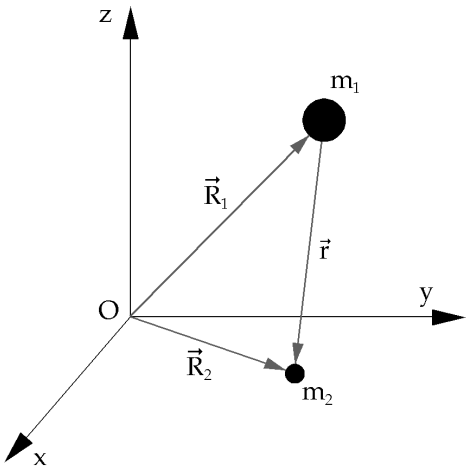

Figure A.3: Geometry of the two bodies problem

From now on the problem considered will be only the two body one, with mass m_1 and m_2 where $m_1 > m_2$; let them be spherical symmetric and the only one force acting upon them would be the gravitational one. Considering an inertial reference frame and being $\vec{\mathbf{R}}_1$ and $\vec{\mathbf{R}}_2$ the position vectors of the two bodies respect to that reference system (cf. fig. A.3), the position of m_2 respect to m_1 will be given by:

$$\vec{\mathbf{r}} = \vec{\mathbf{R}}_2 - \vec{\mathbf{R}}_1$$

taking the derivative one can obtain relative velocity and acceleration:

$$\dot{\vec{\mathbf{r}}} = \dot{\vec{\mathbf{R}}}_2 - \dot{\vec{\mathbf{R}}}_1; \;\; \ddot{\vec{\mathbf{r}}} = \ddot{\vec{\mathbf{R}}}_2 - \ddot{\vec{\mathbf{R}}}_1$$

Once one express the force acting upon the two masses, by means of the second Newton law it can be written the equation of motion. Adding one the other and collecting the terms the following result is obtained:

$$\ddot{\vec{\mathbf{r}}} = -G\frac{m_1 + m_2}{r^3}\vec{\mathbf{r}}$$

Usually can be assumed the ratio between the two masses $m_2/m_1 = m/M$ is much smaller than one[1]. Defining then $\mu = GM$, called gravitational parameter of the primary body, one obtain the Equation of Motion:

$$\ddot{\vec{\mathbf{r}}} = -\frac{\mu}{r^3}\vec{\mathbf{r}} \tag{A.1}$$

It would be observed how that equation does not depend on the mass m, for $m_2 \ll m_1$ the problem become the so called restricted two body problem.

A.1.1 Costant of motion

Starting from the equation of motion A.1 and taking the dot product with the velocity vector $\dot{\vec{\mathbf{r}}}$ and the cross product with the position vector $\vec{\mathbf{r}}$ one can see that the quantities

$$\mathcal{E} = \frac{v^2}{2} - \frac{\mu}{r} \tag{A.2}$$

called mechanic energy and

$$\vec{\mathbf{h}} = \vec{\mathbf{r}} \times \vec{\mathbf{v}} \tag{A.3}$$

named angular momentum, are constant. Moreover since the vector $\vec{\mathbf{h}}$ is normal respect to the orbital plane this means that the motion will take place always on the same plane. Finally computing the cross product between the equation of motion and the angular momentum one obtain another constant of motion, the eccentricity:

$$\vec{\mathbf{e}} = \frac{\vec{\mathbf{v}} \times \vec{\mathbf{h}}}{\mu} - \frac{\vec{\mathbf{r}}}{r}$$

Trajectory Equation

In order to derive the trajectory equation one can dot multiply the position vector $\vec{\mathbf{r}}$ with the eccentricity vector $\vec{\mathbf{e}}$, it results in:

$$\vec{\mathbf{r}} \cdot \vec{\mathbf{e}} = \frac{h^2}{\mu} - r \tag{A.4}$$

[1]The two celestial bodies better balanced in the solar system are Pluto and its satellite Caronte, for them this value is 0.142857.

then letting ν be the angle between the eccentricity vector and the position one it is obtained $\vec{\mathbf{r}} \cdot \vec{\mathbf{e}} = re\cos\nu$ from which it can be derived the trajectory equation in polar coordinates centered in the primary body, that of mass M:

$$r = \frac{h^2/\mu}{1 + e\cos\nu} \tag{A.5}$$

A.1.2 Conic Sections

It is called conic section the set of points P for which the ratio between the distance r from one fixed point F (foci), and the distance d from one given line a is a positive constant e (eccentricity): $r/d = e$. Named s the distance between the foci and the line one has:

$$d = s - r\cos\nu \Rightarrow r = e\,(s - r\cos\nu)$$

Collecting the radius r in the previous equation it becames:

$$r = \frac{es}{1 + e\cos\nu}$$

This is the generic equation for a generic conic section in polar coordinate with the foci at the origin. The constant $p = se$ is called semi-latus rectum or parameter and it is the distance between the point P and the foci for $\nu = \pi/2$. Comparing the previous formula for the trajectory equation with that for a generic conic section it is clear how in the two bodies problem the motion of the smaller mass it is a generic conic and the parameter of the orbit is directly linked to the angular momentum $p = h^2/\mu$.

Relation between mechanic energy and semimajor axis

It is easy to obtain a direct relation between the mechanic energy $\mathcal{E}$ and the semimajor axis of the orbit described by the mass m simply considering the situation at the periaxis:

$$\mathcal{E} = -\frac{\mu}{2a} \tag{A.6}$$

Ellipse

An ellipse is a particular conic section that can be described as the set of point for which the sum of the distance from two fixed points (focii) is constant:

$$r_1 + r_2 = 2a$$

The biggest dimension is called semimajor axis and measure a, the distance drawed perpendicularly to the semimajor axis crossing the center of the curve

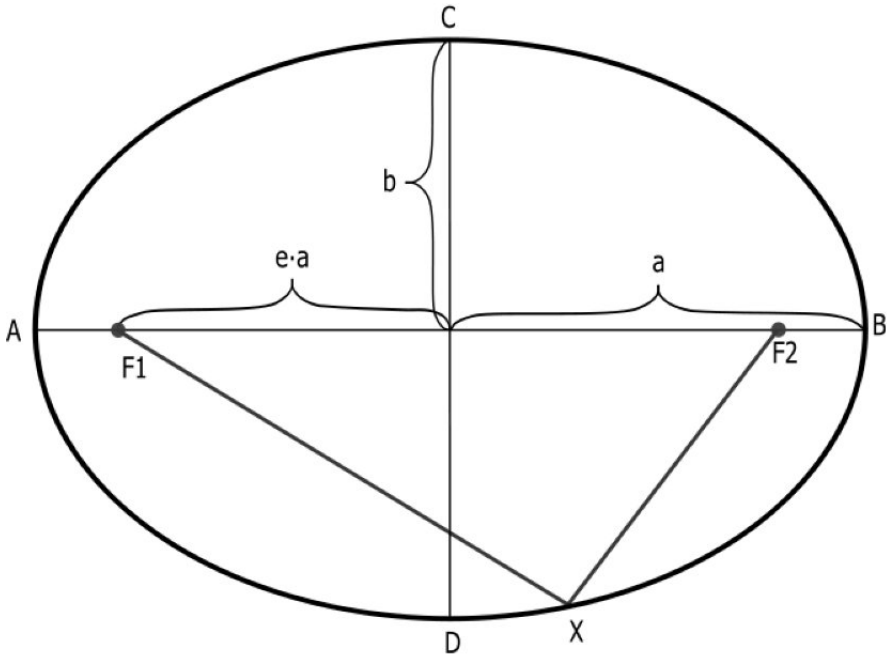

Figure A.4: Geometry of an Ellipse

is named semiminor axis and measure $2b$. The distance between the focii instead is $2c$.

The following relations can be derived from the observation of the figure A.4:

$$\begin{aligned} r_P &= p/(1+e) && \text{(A.7)} \\ 2a &= r_P + r_A = 2p/\left(1-e^2\right) && \text{(A.8)} \\ r_A &= p/(1-e) && \text{(A.9)} \\ 2c &= r_A - r_P = 2pe/\left(1-e^2\right) = 2ae && \text{(A.10)} \\ b^2 &= a^2 - c^2 = a^2\left(1-e^2\right) = ap && \text{(A.11)} \end{aligned}$$

The area can be expressed as:

$$A = \pi ab = \pi a^2\sqrt{(1-e^2)} = \pi a\sqrt{ap}$$

A.2 Two dimensional motion analysis

In order to study the motion of one body in the space it is necessary to take always present all the conventions used and with them all the parameters

that will be adopted in order to identify the position in space.

Reference systems and velocity and acceleration components

The reference system choice is the starting point for a motion analysis. As already seen in the two bodies problem the orbit plane remains always the same, for this reason the motion analysis become bidimensional. Considering the constance of the direction and magnitude of the angular momentum A.3, one can define the various versors of the orthogonal right hand ruled non rotating reference system with the origin in M as:

$$\hat{\mathbf{n}} = \text{chosen arbitrarily on the orbit plane} \quad \hat{\mathbf{m}} = \frac{\vec{\mathbf{h}} \times \hat{\mathbf{n}}}{h}; \quad \hat{\mathbf{k}} = \frac{\vec{\mathbf{h}}}{h}$$

Then can be introduced another reference system, this time not fixed but still right hand ruled and orthogonal, with the origin on the mass m and with versors defined as:

$$\hat{\mathbf{i}} = \frac{\vec{\mathbf{r}}}{r}; \quad \hat{\mathbf{j}} = \frac{\vec{\mathbf{h}} \times \hat{\mathbf{i}}}{h}; \quad \hat{\mathbf{k}} = \frac{\vec{\mathbf{h}}}{h}$$

In this last system, called local reference - $\mathcal{F}_L$, are here defined the velocity vector and acceleration vector components, in this case one must take into account also the temporal derivatives of the versors that are not fixed:

$$\begin{aligned} v_r &= \dot{r} & v_\vartheta &= r\dot{\vartheta} \\ a_r &= \ddot{r} - r\dot{\vartheta}^2 & a_\vartheta &= r\ddot{\vartheta} + 2\dot{r}\dot{\vartheta} \end{aligned}$$

Among the various non rotating systems one possible choiche is that called perifocal reference system - $\mathcal{F}_P$, centered in the body with mass M, and with versors defined as:

$$\hat{\mathbf{p}}_1 = \frac{\vec{\mathbf{e}}}{e}; \quad \hat{\mathbf{p}}_2 = \frac{\vec{\mathbf{h}} \times \hat{\mathbf{p}}_1}{h}; \quad \hat{\mathbf{p}}_3 = \hat{\mathbf{k}} = \frac{\vec{\mathbf{h}}}{h}$$

One should observe that the angle ω between the versor $\vec{\mathbf{p}}_1$ and $\vec{\mathbf{n}}$ previously defined is constant; so for passing from one general fixed reference on the orbital plane to the perifocal one the same relations for velocity and acceleration must hold with the following device:

$$\vartheta = \nu + \omega; \quad \dot{\nu} = \dot{\vartheta}; \quad \ddot{\nu} = \ddot{\vartheta}$$

The three systems rappresentation is drawed in the following figure A.5.

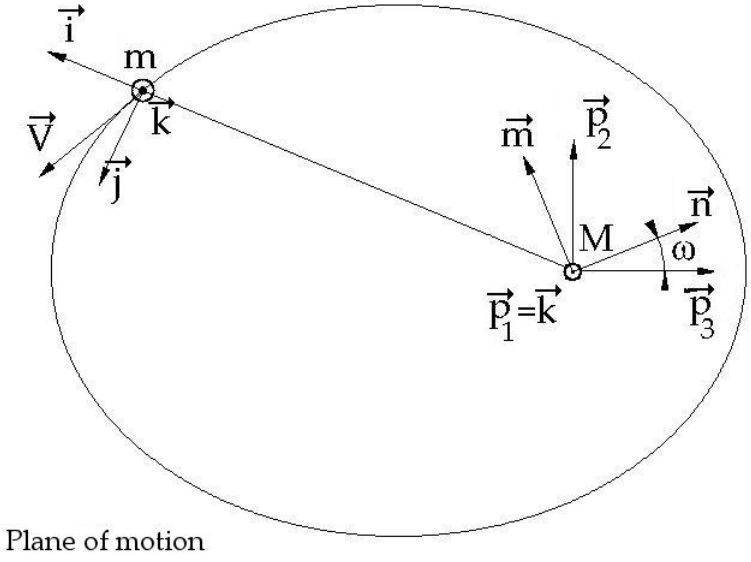

Figure A.5: Reference Systems

Flight Path Angle

Let is consider the situation rappresented in figure A.6; let s be the generic trajectory of the mass m and F the position of the mass M. In order to evaluate the velocity vector components v_r and v_ϑ that the mass m has in a generic point on the orbit respect to the versors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ of the local reference it must be considered the generic trajectory equation

$$r\left(1+e\cos\nu\right)=p=cost$$

taking the time derivative

$$\dot{r}\left(1+e\cos\nu\right)-re\dot{\nu}\sin\nu=0$$

so

$$\tan\phi=\frac{v_r}{v_\vartheta}=\frac{\dot{r}}{r\dot{\nu}}=\frac{e\sin\nu}{1+e\cos\nu} \tag{A.12}$$

where the angle ϕ is the angle between the velocity $\vec{\mathbf{V}}$ and the versor $\hat{\mathbf{j}}$, perpendicular to the local vertical $\hat{\mathbf{i}}$, so which lie in the orizzontal local plane, this angle is named flight path angle.

A.2.1 Time of flight on the elliptic orbit

Starting from the analysis of the elementary piece of area sweeped by the position vector of the mass m one can obtain the expression of the period for

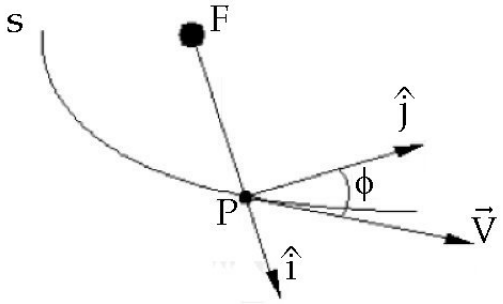

Figure A.6: Flight Path Angle

an elliptic orbit as follows:

$$T = 2\pi\sqrt{\frac{a^3}{\mu}}$$

In order to evaluate the time of flight of an elliptic orbit arc it must be considered the area enclosed by the position vector of the mass m in the initial and final position, then there is a proportional relation between this area A, the total ellipse area πab, the orbit period T and the time of flight from the starting to the ending point on the orbit $t_{1\to 2}$. Assuming as initial position the periaxis, the time of flight from it to a generic point P along the orbit can be expressed as:

$$t_P = \sqrt{\frac{a^3}{\mu}}\,(E - e\sin E)$$

where E is named eccentric anomaly (cf. fig. A.7), and it is linked to the ellipse parameters by means of the following formulas:

$$\cos E = \frac{a - r}{ea} = \frac{e + \cos\nu}{1 + e\cos\nu} \tag{A.13}$$

For evaluate the time of flight from a generic point P_1 to another generic point P_2 one can write:

$$t_{1\to 2} = t_{P_2} - t_{P_1} = \sqrt{\frac{a^3}{\mu}}[E_2 - E_1 - e\,(\sin E_2 - \sin E_1)] \tag{A.14}$$

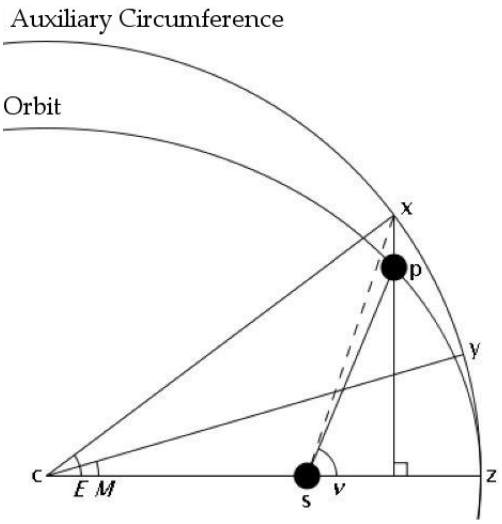

Figure A.7: Eccentric Anomaly

Extension to Parabolic and Hyperbolic Orbits

It is possible to define in a similar way the time of flight along non elliptic orbit arches. In case of hyperbolic orbits the relation will be:

$$t_{1\to 2} = t_{P_2} - t_{P_1} = \sqrt{-\frac{a^3}{\mu}}[e\,(\sinh F_2 - \sinh F_1) - (F_2 - F_1)] \tag{A.15}$$

where F is named hyperbolic eccentric anomaly and is linked to the trajectory parameters by means of the quantities:

$$\cosh F = \frac{a - r}{ea} = \frac{e + \cos\nu}{1 + e\cos\nu} \tag{A.16}$$

In case of parabolic orbits one has

$$t_P = \frac{1}{2\sqrt{\mu}}\left(pD + \frac{1}{3}D^3\right) \tag{A.17}$$

where D is the parabolic eccentric anomaly, defined as

$$D = \sqrt{p}\tan\left(\frac{\nu}{2}\right)$$

Circular Velocity

With the term circular velocity is indicated that velocity which would allow the mass m to describe a circular orbit with a null radial velocity $v_r = 0$.

This quantity can be obtained equalizing the centrifugal force, caused by the centripetal acceleration, to the gravitational attraction force, it is value is:

$$v_c = \sqrt{\frac{\mu}{r}} \tag{A.18}$$

A.3 Three dimensional motion analysis

A.3.1 Orbital Parameters

Classical Orbital Parameters

A Keplerian trajectory is univocally determined by six parameter, since it is the solution of a three dimensional vectorial second–order differential equation; the classical orbital parameters are widely used with this aim: four of them are necessary for the bidimensional analysis indentifying the trajectory on the orbital plane, while in order to place this plane in space two more parameters are needed.

These parameteres are (cf. fig. A.8):

- a semimajor axis or p parameter - (orbit dimension);
- e eccentricity - (orbit shape);
- i inclination - (inclination of the orbital plane respect to the reference plane);
- Ω longitude of the ascending node - right ascension of the ascending node (RAAN) - (angular distance mesured on the reference plane between the point in where the orbit crosses the reference plane while moving towards the north and a fixed axis previously wisely chosen);
- ω argument of the periaxis - (angular distance mesured on the orbital plane between the periaxis and the ascending node);
- ν_0 true anomaly at epoch t_0 or M_0 mean anomaly at epoch t_0 - (position of the mass m in a fixed time);

Modified Equinoctial Parameters

The previous set of orbital parameters has some problem in cas of orbits of zero inclination (the RAAN is not yet defined) and if the orbits are circular (so the argument of periapsis become meaningless); a possible solution

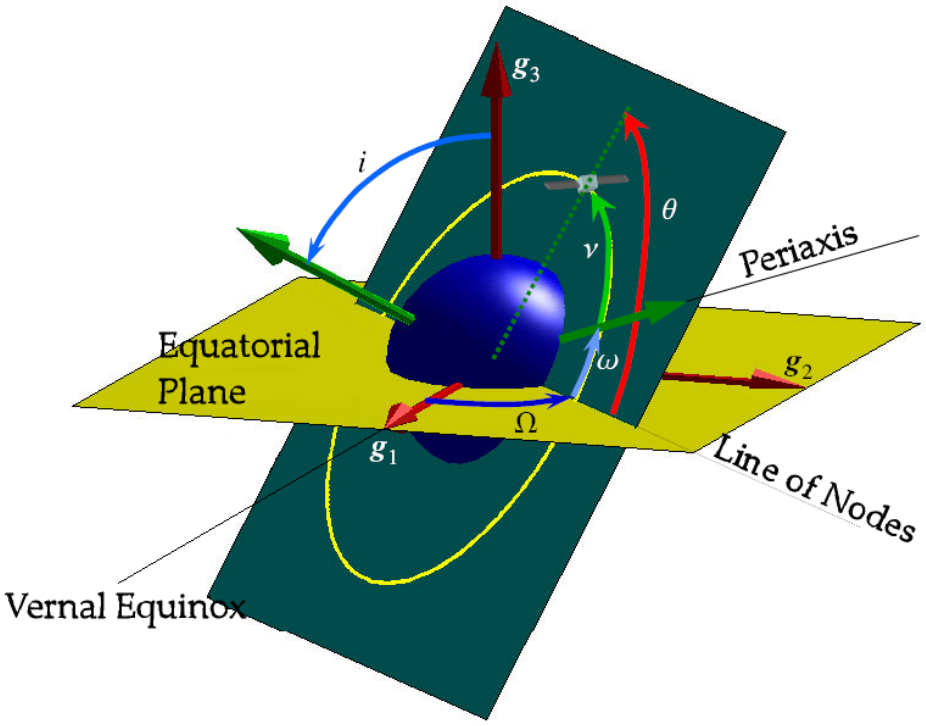

Figure A.8: Classical Orbital Parameters

is rappresented by the so called Modified Equinoctial Paramters, these are always defined for all eccentricity and inclination between 0° and 180°, so the retrograde orbits are excluded but this is usually not a problem; these parameters are

- $p = a\,(1 - e^2)$;
- $f = e\cos(\omega + \Omega)$;
- $g = e\sin(\omega + \Omega)$;
- $h = \tan(i/2)\cos\Omega$;
- $k = \tan(i/2)\sin\Omega$;
- $l_0 = \Omega + \omega + \nu_0$ or $l_0 = \Omega + \omega + M_0$;

computing the inverse relations one has

- $a = \frac{p}{1-f^2-g^2}$;

- $e = \sqrt{f^2 + g^2}$;
- $i = 2\tan^{-1}(h^2 + k^2)$;
- $\omega = \tan^{-1}(g/f) - \tan^{-1}(k/h)$;
- $\Omega = \tan^{-1}(k, h)$;
- $\nu_0 = l_0 - \Omega - \omega$;

Determining the orbital elements

The orbital elements can be easily calculated starting from the cartesian components of $\vec{\mathbf{r}}$ and $\vec{\mathbf{v}}$ measured in an intertial reference system at a particular time t_0. First of all one has to rememeber the relations

$$\vec{\boldsymbol{h}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}}; \quad \vec{\boldsymbol{e}} = \frac{\vec{\boldsymbol{v}} \times \vec{\boldsymbol{h}}}{\mu} - \frac{\vec{\boldsymbol{r}}}{r}$$

and so it can be defined the unit vectors as follows

$$\hat{\boldsymbol{k}} = \frac{\vec{\boldsymbol{h}}}{h}; \quad \hat{\boldsymbol{i}} = \frac{\vec{\boldsymbol{r}}}{r} \quad \hat{\boldsymbol{n}} = \frac{\vec{\boldsymbol{g}}_3 \times \vec{\boldsymbol{k}}}{\left\|\vec{\boldsymbol{g}}_3 \times \vec{\boldsymbol{k}}\right\|} \quad \hat{\boldsymbol{p}}_1 = \frac{\vec{\boldsymbol{e}}}{e}$$

these unit vectors centered in the primary body define the direction normal to the orbit plane, the direction towards the smaller mass m, the direction of the ascending node and the direction of the periapsis, respectively. The orbital elements can be computed as

- $p = \frac{h^2}{\mu}$;
- $e = \|\vec{\boldsymbol{e}}\|$;
- $\cos i = \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{g}}_3 = k_3$;
- $\cos\Omega = \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{g}}_1 = n_3$ $(n_2 < 0 \Rightarrow \Omega > \pi)$;
- $\cos\omega = \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{p}}_1$ $(e_3 < 0 \Rightarrow \omega > \pi)$;
- $\cos\nu_0 = \hat{\boldsymbol{i}} \cdot \hat{\boldsymbol{p}}_1$ $(\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{v}} < 0 \Rightarrow \nu_0 > \pi)$;

Expressing $\vec{\mathbf{r}}$ and $\vec{\mathbf{v}}$ in the Perifocal System

Supposing to know all the orbital parameters, it is simple to express the position in the perifocal system

$$\vec{\boldsymbol{r}} = r\cos\nu\hat{\boldsymbol{P}} + r\sin\nu\hat{\boldsymbol{Q}}$$

where for the magnitude of r one can use the expression in polar coordinates

$$r = \frac{r}{1 + e\cos\nu}$$

In order to obtain $\vec{\mathbf{v}}$ one can differentiate the previous formula with the derivatives of the unit vectors equal to zero since the frame is inertial, so

$$\vec{v} = \dot{\vec{\boldsymbol{r}}} = \left(\dot{r}\cos\nu - r\dot{\nu}\sin\nu\right)\hat{\boldsymbol{P}} + \left(\dot{r}\sin\nu + r\dot{\nu}\cos\nu\right)\hat{\boldsymbol{Q}}$$

that can be simplified noting that

$$\dot{r} = \sqrt{\tfrac{\mu}{p}}e\sin\nu \quad r\dot{\nu} = \sqrt{\tfrac{\mu}{p}}\left(1 + e\cos\nu\right)$$

and it becomes

$$\vec{\boldsymbol{v}} = \sqrt{\frac{\mu}{p}}\left[-\sin\nu\hat{\boldsymbol{P}} + \left(e + \cos\nu\right)\hat{\boldsymbol{Q}}\right]$$

A.4 Perturbations

Since now the discussion has involved only the motion of a body under the influence of the gravitational force, the reality is far from this situation: all the other forces (like the presence of the other bodies, the atmospheric drag, the solar pressure, etc.) can be considered as perturbations. A perturbation is a deviation from the normal or expected behaviour; some of these perturbations maybe unpredictable and so must be treated in a stochastic/probabilistic way, other ones can be modelled and analyzed analitically; sometimes these deviation can be as large as the primary force so these added forces must be always taken into account. There are two principal perturbation techniques: the first one, special perturbation, consist of the numerical integration of motion including all the perturbing acceleration and the second one, general perturbation, consists in an analytical integration of series expansions of the perturbing acceleration.

A.4.1 Encke's Method

The Encke's Method is one of the most famous special perturbation methods. The principle which lay in the background is to integrate separately the two accelerations, this implies the existence of a reference orbit along which the object would move under the influence of only the primary force, it means the trajectory would be a conic section. All the calculations would be respect to this orbit which is called osculating orbit since it is the kissing keplerian orbit at a determined epoch, and would be the orbit since that epoch going forward if all the perturbations are removed (cf. A.9). One can consider a keplerian

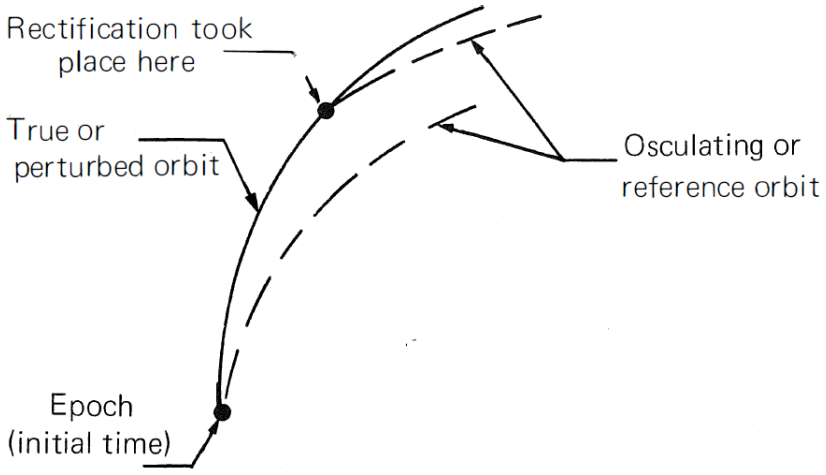

Figure A.9: Osculating orbit with retification

(kissing) orbit that is good until the perturbations does not deviate it too far, at this time a rectification process must occur to continue the integration, this means defining as osculating orbit the newer (true) one derived from the integration at a new point. Considering the analytical formulation: first of all it must be find an analytic expression of the difference between the reference orbit and the true one, let $\vec{\boldsymbol{r}}$ and $\vec{\boldsymbol{\rho}}$ be the radius vector of, respectively, the true orbit and the osculating one at a particular time. So

$$\ddot{\vec{\boldsymbol{r}}} + \frac{\mu}{r^3}\vec{\boldsymbol{r}} = \vec{\boldsymbol{a}}_P$$

and

$$\ddot{\vec{\boldsymbol{\rho}}} + \frac{\mu}{\rho^3}\vec{\boldsymbol{\rho}} = 0$$

since at the epoch $t_0 = 0$ is

$$\vec{\boldsymbol{r}}\,(t_0) = \vec{\boldsymbol{\rho}}\,(t_0) \text{ and } \vec{\boldsymbol{v}}\,(t_0) = \dot{\vec{\boldsymbol{\rho}}}\,(t_0)$$

The deviation from the reference orbit can be defined as (cf. A.10)

$$\vec{\boldsymbol{\delta r}} = \vec{\boldsymbol{r}} - \vec{\boldsymbol{\rho}}$$

so

$$\ddot{\vec{\boldsymbol{\delta r}}} = \ddot{\vec{\boldsymbol{r}}} - \ddot{\vec{\boldsymbol{\rho}}}$$

Substituting the previous equation one has

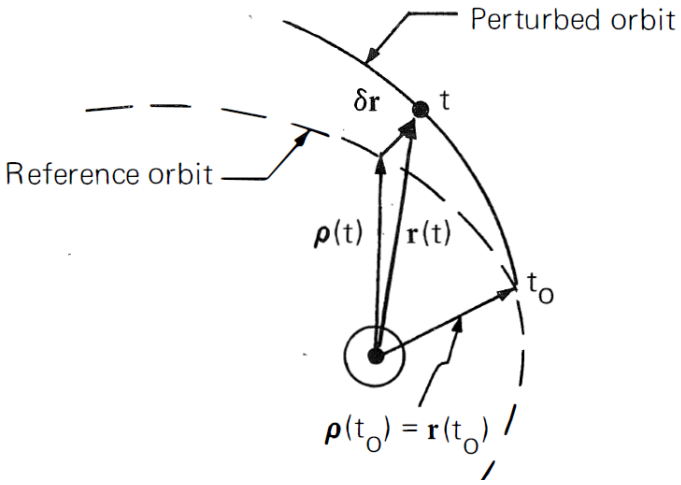

Figure A.10: δr deviation from the reference orbit

$$\ddot{\vec{\boldsymbol{\delta r}}} = \vec{\boldsymbol{a}}_P + \frac{\mu}{\rho^3}\left[\left(1 - \frac{\rho^3}{r^3}\right)\vec{\boldsymbol{r}} - \vec{\boldsymbol{\delta r}}\right]$$

so one has the differential equation for the deviation from the reference orbit and δr can be evaluated numerically, since $\vec{\boldsymbol{\rho}}$ is know one can derive $\vec{\boldsymbol{r}}$ from the previous two qunatities. With that precedure it is possible to evaluate every true position forward in time. Anyway there is a problem since the term

$$1 - \frac{\rho^3}{r^3}$$

is the difference of two very nearly equal quantities and require very little steps of integration in order to obtain a reasonable accuracy. This problem can be faced defining

$$2q = 1 - \frac{r^2}{\rho^2}$$

so

$$\frac{\rho^3}{r^3} = (1 - 2q)^{-3/2}$$

and the previous equation become

$$\ddot{\vec{\boldsymbol{\delta}}}\boldsymbol{r} = \vec{\boldsymbol{a}}_P + \frac{\mu}{\rho^3}\left\{\left[1-(1-2q)^{-3/2}\right]\vec{\boldsymbol{r}} - \vec{\boldsymbol{\delta}}\boldsymbol{r}\right\}$$

Now it must be founded an expression for the quantity q, starting from the definition of r one can obtain

$$q = -\frac{1}{\rho^2}\left[\delta x\left(\rho_x + \frac{1}{2}\delta x\right) + \delta y\left(\rho_y + \frac{1}{2}\delta y\right) + \delta z\left(\rho_z + \frac{1}{2}\delta z\right)\right]$$

and the expansion in series of q

$$1-(1-2q)^{-3/2} = -3q - \frac{15}{2!}q^2 - \frac{105}{3!}q^3 - ...$$

The method can be implemented as follows

1. Initial condition known: $\vec{\boldsymbol{r}}(t_0) = \vec{\boldsymbol{\rho}}(t_0)$, $\vec{\boldsymbol{v}}(t_0) = \dot{\vec{\boldsymbol{\rho}}}(t_0)$ defining the osculating orbit; also at this point one has $\vec{\boldsymbol{\delta}}\boldsymbol{r} = \delta r = 0$;
2. Calculate for one integration step $\delta r(t_0 + \Delta t)$ knowing $\vec{\boldsymbol{\rho}}(t_0)$, $r(t_0)$; $q(t_0) = 0$;
3. Knowing $\vec{\boldsymbol{\delta}}\boldsymbol{r}(t_0 + \Delta t)$ one can compute $\vec{\boldsymbol{\rho}}(t_0 + \Delta t)$ and $q(t_0 + \Delta t)$;
4. Integrate another Δt to get $\vec{\boldsymbol{\delta}}\boldsymbol{r}(t + k\Delta t)$;
5. If $\frac{\delta r}{\rho} >$ a fixed value go to the first step otherwise continue;
6. Calculate $\vec{\boldsymbol{r}} = \vec{\boldsymbol{\rho}} + \vec{\boldsymbol{\delta}}\boldsymbol{r}$ and $\vec{\boldsymbol{v}} = \dot{\vec{\boldsymbol{\rho}}} + \dot{\vec{\boldsymbol{\delta}}}\boldsymbol{r}$;
7. Go to step three where Δt is replaced by $k\Delta t$ where k is the step number;

A.4.2 Gauss Equations

The method of the variation of parameters was originally developed by Lagrange in order to study the perturbed motion of two bodies in the form

$$\frac{d\vec{\boldsymbol{r}}}{dt} = \vec{\boldsymbol{v}} \quad \frac{d\vec{\boldsymbol{v}}}{dt} + \frac{\mu}{r^3}\vec{\boldsymbol{r}} = \left[\frac{\partial R}{\partial \vec{\boldsymbol{r}}}\right]^T \tag{A.19}$$

where R is the disturbing function defined as

`Section 8.4 Battin`

The solution of the undisturbed motion is known, can be expressed as

$$\vec{\boldsymbol{r}} = \vec{\boldsymbol{r}}\,(t, \vec{\boldsymbol{\alpha}}) \quad \vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}\,(t, \vec{\boldsymbol{\alpha}})$$

where the components of the vector $\vec{\boldsymbol{\alpha}}$ are the six constants of motion, e.g. orbital elements. One can assume that under the perturbations the parameters $\vec{\boldsymbol{\alpha}}$ can vary in time, in order to satisfy the equation A.19 for the disturbed motion. With this observation a set of differential equations for α can be derived, these are in fact the result of the trasformation of the dependent variables of the problem from the position and velocity vector to the parameters; although these new equations are complex as the previous ones they have many advantages since only the perturbing acceleration will affect the change in time of the variables so is a sort of generalization of the Encke's Method in a continuous one. The derivation need the following steps

$$\tfrac{\partial \vec{\boldsymbol{r}}}{\partial t} = \vec{\boldsymbol{v}} \quad \tfrac{\partial \vec{\boldsymbol{v}}}{\partial t} + \tfrac{\mu}{r^3}\vec{\boldsymbol{r}} = 0$$

for the disturbed motion

$$\frac{d\vec{\boldsymbol{r}}}{dt} = \frac{\partial \vec{\boldsymbol{r}}}{\partial t} + \frac{\partial \vec{\boldsymbol{r}}}{\partial \alpha}\frac{d\vec{\boldsymbol{\alpha}}}{dt}$$

and comparing this result with the previous equation one obtain

$$\frac{\partial \vec{\boldsymbol{r}}}{\partial \vec{\boldsymbol{\alpha}}}\frac{d\vec{\boldsymbol{\alpha}}}{dt} = 0 \tag{A.20}$$

which means that the velocity vector of the disturbed and undesturbed motion must be identical. Similarly

$$\frac{d\vec{\boldsymbol{v}}}{dt} = \frac{\partial \vec{\boldsymbol{v}}}{\partial t} + \frac{\partial \vec{\boldsymbol{v}}}{\partial \vec{\boldsymbol{\alpha}}}\frac{d\vec{\boldsymbol{\alpha}}}{dt}$$

again comparing the result with the previous equation one obtain

$$\frac{\partial \vec{\boldsymbol{v}}}{\partial \vec{\boldsymbol{\alpha}}}\frac{d\vec{\boldsymbol{\alpha}}}{dt} = \left[\frac{\partial R}{\partial \vec{\boldsymbol{r}}}\right]^T \tag{A.21}$$

and the equations A.20 and A.21 are the six scalar equations for the parameter $\vec{\alpha}$.

Writing the equation for the orbital elements

Choosing the six parameters as

$$\vec{\boldsymbol{\alpha}}^T = \begin{bmatrix} \Omega & i & \omega & a & e & \lambda \end{bmatrix}$$

and compunting all the calculation (see [2]) one can obtain

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{r\sin(\nu+\omega)}{h\sin i}a_w \\
\frac{di}{dt} &= \frac{r\cos(\nu+\omega)}{h}a_w \\
\frac{d\omega}{dt} &= \frac{1}{he}\left[-p\cos\nu a_r + (p+r)\sin\nu a_\theta\right] - \frac{d\Omega}{dt}\cos i \\
\frac{da}{dt} &= \frac{2a^2}{h}\left(e\sin\nu a_r + \frac{p}{r}a_\theta\right) \\
\frac{de}{dt} &= \frac{1}{h}\left\{p\sin\nu a_r + \left[(p+r)\cos\nu + re\right]a_\theta\right\} \\
\frac{dM}{dt} &= n + \frac{b}{ahe}\left[(p\cos\nu - 2re)\,a_r - (p+r)\sin\nu a_\theta\right]
\end{aligned} \tag{A.22}$$

obtaining the Lagrange's Planetary equations. If one want to obtain the equations with the disturbing acceleration components expressed in the reference system with one direction along the velocity vector and the other one normal to it, can observe that

$$\begin{bmatrix} a_r \\ \\ a_\theta \end{bmatrix} = \frac{h}{pv}\begin{bmatrix} e\sin\nu & -(1+e\cos\nu) \\ 1+e\cos\nu & e\sin\nu \end{bmatrix}\begin{bmatrix} a_t \\ \\ a_n \end{bmatrix}$$

substituting in the previous system one obtain

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{r\sin(\nu+\omega)}{h\sin i}a_w \\
\frac{di}{dt} &= \frac{r\cos(\nu+\omega)}{h}a_w \\
\frac{d\omega}{dt} &= \frac{1}{ev}\left[2\sin\nu a_t + \left(2e + \frac{r}{a}\cos\nu\right)a_n\right] - \frac{d\Omega}{dt}\cos i \\
\frac{da}{dt} &= \frac{2a^2v}{\mu}a_t \\
\frac{de}{dt} &= \frac{1}{v}\left[2\left(e+\cos\nu\right)a_t - \frac{r}{a}\sin\nu a_n\right] \\
\frac{dM}{dt} &= n - \frac{b}{eav}\left[2\left(1+\frac{e^2r}{p}\right)\sin\nu a_t + \frac{r}{a}\cos\nu a_n\right]
\end{aligned} \tag{A.23}$$

Writing the equation for nonsingular elements

The previous system of equations has some critical problems, infact for zero inclination angle or for orbit of zero eccentricity the equations became singular. In order to solve these problems one can use the so called modified equinocial elements

$$\begin{aligned}
P_1 &= e\sin\varpi \quad Q_1 = \tan\tfrac{1}{2}i\sin\Omega \\
P_2 &= e\cos\varpi \quad Q_1 = \tan\tfrac{1}{2}i\cos\Omega
\end{aligned}$$

where

$$\varpi = \Omega + \omega \quad l = \varpi + M \quad L = \varpi + \nu$$

in this way one obtains

$$\begin{aligned}
&\tfrac{da}{dt} = \tfrac{2a^2}{h}\left[\left(P_2 \sin L - P_1 \cos L\right) a_r + \tfrac{p}{r} a_\theta\right] \\
&\tfrac{dP_1}{dt} = \tfrac{r}{h}\left\{-\tfrac{p}{r}\cos L a_r + \left[P_1 + \left(1 + \tfrac{p}{r}\right)\sin L\right] a_\theta - P_2\left(Q_1 \cos L - Q_2 \sin L\right) a_w\right\} \\
&\tfrac{dP_2}{dt} = \tfrac{r}{h}\left\{\tfrac{p}{r}\sin L a_r + \left[P_2 + \left(1 + \tfrac{p}{r}\right)\cos L\right] a_\theta + P_1\left(Q_1 \cos L - Q_2 \sin L\right) a_w\right\} \\
&\tfrac{dQ_1}{dt} = \tfrac{r}{2h}\left(1 + Q_1^2 + Q_2^2\right)\sin L a_w \\
&\tfrac{dQ_2}{dt} = \tfrac{r}{2h}\left(1 + Q_1^2 + Q_2^2\right)\cos L a_w \\
&\tfrac{dl}{dt} = n - \tfrac{r}{h}\left\{\left[\tfrac{a}{a+b}\left(\tfrac{p}{r}\right)\left(P_1 \sin L + P_2 \cos L\right) + \right.\right. \\
&\left.\left. + \tfrac{2b}{a}\right] a_r + \tfrac{a}{a+b}\left(1 + \tfrac{p}{r}\right)\left(P_1 \cos L - P_2 \sin L\right) a_\theta + \left(Q_1 \cos L - Q_2 \sin L\right) a_w\right\}
\end{aligned}$$

which are equations with no singularities.

Appendix B

Fundamentals of Perturbation Methods

In this chapter there will be outlined the main characteristics of perturbation methods, starting from general definitions to the standard and singular perturbation methods, with some example made for a better understanding of their possible applications.

B.1 Asimptotic Expansion

Consider a sequence of functions $\{\phi_n(\varepsilon)\}$, $n = 1, 2, ...$ Such a sequence is called an asymptotic sequence if

$$\phi_{n+1} = o(\phi_n(\varepsilon)) \text{ as } \varepsilon \to 0$$

for each $n = 1, 2, ...$

Notice that the previous definition does not preclude having one or more of the starting terms in an asymptotic sequence being infinite. Here again, various operations, such as multiplication of two sequences or integration, can be used to generate a new sequence. Differentiation with respect to e may not lead to a new asymptotic sequence. Let $u(\boldsymbol{x};\varepsilon)$ be defined in some domain D of $\boldsymbol{x}$ and some neighborhood of $\varepsilon = 0$. Let $\{\phi_n(\varepsilon)\}$ be a given asymptotic sequence. The series $\sum_{n=1}^{M} \phi_n(\varepsilon)\, u(\boldsymbol{x})$ is called the asymptotic expansion of $u(\boldsymbol{x};\varepsilon)$ to N terms (N may be a finite integer or infinity) as $\varepsilon \to 0$ with respect to the sequence $\phi_n(\varepsilon)$ if

$$u(\boldsymbol{x};\varepsilon) - \sum_{n=1}^{M} \phi_n(\varepsilon)\, u(\boldsymbol{x}) = o(\phi_M) \text{ as } \varepsilon \to 0 \tag{B.1}$$

for each $M = 1, 2, ..., N$. And if the relation B.1 holds uniformly in all the interval D the expansion in said to be uniformly valid.

Once the function $u(\boldsymbol{x}; \varepsilon)$ is given and the asymptotic sequence $\phi_n()$ is specified, it can be defined each of the $u_n(\boldsymbol{x})$ uniquely by repeated application of definition

$$u_1(\boldsymbol{x}) = \lim_{\varepsilon \to 0} \frac{u(\boldsymbol{x}; \varepsilon)}{\phi_1(\varepsilon)} \tag{B.2}$$

$$u_k(\boldsymbol{x}) = \lim_{\varepsilon \to 0} \frac{u(\boldsymbol{x}; \varepsilon) - \sum_{n=1}^{k-1} \phi_n(\varepsilon) u_n(\boldsymbol{x})}{\phi_k(\varepsilon)} \tag{B.3}$$

This approach must be applied with some advices due to it is particular validity cases, in fact if one consider the next function

$$u(x; \varepsilon) = e^{-\frac{x}{\varepsilon}} - \frac{\varepsilon e^{-x}}{x + \varepsilon} \equiv f(x; \varepsilon) \tag{B.4}$$

in the interval $0 < x < 1$ and $0 < \varepsilon \ll 1$. Fixing it and applying the expansion process with $\phi_n = \varepsilon^{n-1}$ the following result will be obtained

$$f(x; \varepsilon) = -\varepsilon \frac{e^{-x}}{x} + \varepsilon^2 \frac{e^{-x}}{x^2} - \varepsilon^3 \frac{e^{-x}}{x^3} + \mathcal{O}\left(\varepsilon^4\right) = \sum_{n=1}^{N} \varepsilon^n h_n(x) + \mathcal{O}\left(\varepsilon^{N+1}\right) \tag{B.5}$$

This expansion is not uniformly valid in any sub interval with $x = 0$ as

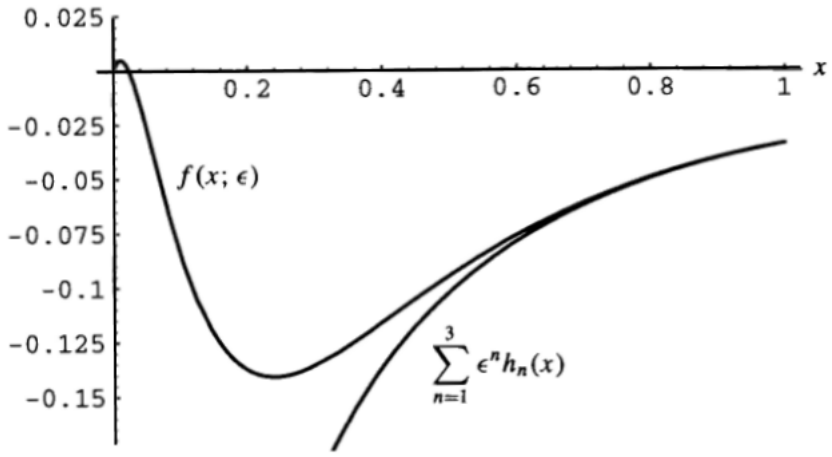

Figure B.1: Exact Solution and Outer Expansion

left limit point, in fact, this expansion is singular at $x = 0$, and it does not provide a good approximation of B.4 no matter how small ε is chosen if x

is allowed to become arbitrarily small. This is seen in fig. B.1, where B.4 is compared with B.5 for the choice $\varepsilon = 0.1$. However, if x is restricted to lie in the subinterval $0 < xo < x < 1$, then B.5 is uniformly valid there. The source of the nonuniformity nearx $x = 0$ is easily traced to the expansion of the denominator $x + \varepsilon$ in the second term. This expansion is based on the limit $\varepsilon \to 0$ with x fixed and is thus incorrect if $x = \mathcal{O}(\varepsilon)$ or smaller. It is natural to seek another expansion that adequately approximates B.4, for x small. Since the nonuniformity occurs for $x = \mathcal{O}(\varepsilon)$, and since the combination x/e appears in the first term of B.4, one is led to the change of variable $x^* = x/e$

$$u = e^{-\frac{x^*}{\varepsilon}} - \frac{\varepsilon e^{-x^*}}{x^* + 1} \equiv g(x^*; \varepsilon) \tag{B.6}$$

in this f and g define the same function with a change in variable but the expansion of the g function is quite different from that for the f, it is

$$\begin{array}{rcl} g(x;\varepsilon) & = & e^{-x^*} - \frac{1}{1+x^*} + \frac{\varepsilon x^*}{1+x^*} - \frac{\varepsilon^2 x^{*2}}{2(1+x^*)} + \frac{\varepsilon^3 x^{*3}}{6(1+x^*)} + \mathcal{O}(\varepsilon^4) = \\ & = & \sum_{n=1}^{N} \varepsilon^{n-1} g_n(x^*) + \mathcal{O}\left(\varepsilon^{N+1}\right) \end{array} \tag{B.7}$$

One can note that this expansion gives a good approximation for u for small x. In particular, the right-hand side of B.7 vanishes at $x^* = x = 0$ to all orders in ε, and this conforms with the exact value of u at $x = 0$. However, B.7 fails to be uniformly valid for $x^* \to \infty$, as seen in fig. B.2. Therefore, the two expansions B.5 and B.7 have mutually exclusive domains of validity. Depending on the magnitude of x compared with ε, one expansion or the other should be used. B.5 is named as outer expansion of u (because it is valid away from the boundary point $x = 0$), and the expansion B.7 is called the inner expansion of u (because it is valid near $x = 0$). It can be found that these two different expansions will produce the same result after the substitution of the variable relative to the other expansion, it means that a matching condition between them can be founded easily.

These kind of problems are called singular due to the presence of the point of singularity.

B.2 Regular Expansions

B.2.1 Regular Expansions for Algebraic Equations

Consider a function $u = f(\boldsymbol{x}; \varepsilon)$ which is defined implicitly as the root of a certain algebraic equation $R(\boldsymbol{x}, u; \varepsilon) = 0$ that cannot be solved explicitly for $f(\boldsymbol{x}; \varepsilon)$ for arbitrary $\varepsilon \neq 0$. If one is interested only in the solution of

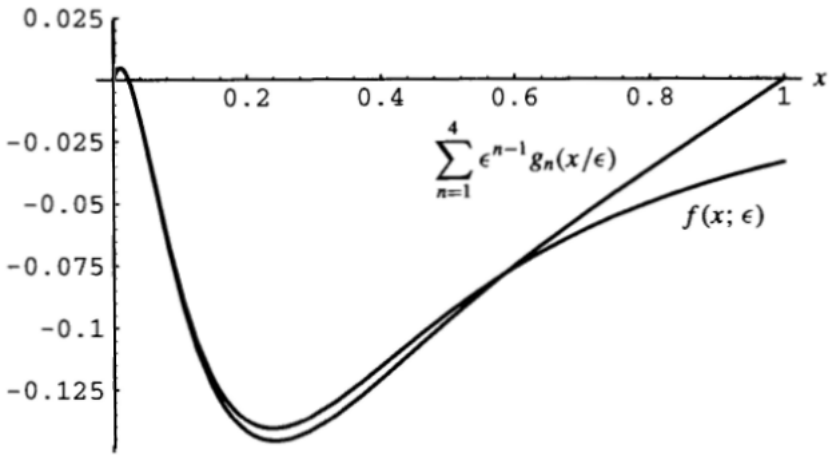

Figure B.2: Exact Solution and Outer Expansion

$R = 0$ for ε small, and if $R(\boldsymbol{x}, u; 0) = 0$ is solvable, it is useful to construct the asymptotic expansion of $f(x;\varepsilon)$ as $\varepsilon \to 0$.

One can create an expansion for f as

$$f(\boldsymbol{x};\varepsilon) \equiv f(\boldsymbol{x};0) + \phi_2(\varepsilon) f_2(\boldsymbol{x}) + o(\phi_2(\varepsilon)) \text{ as } \varepsilon \to 0 \tag{B.8}$$

where ϕ_2 is not given and should be choosen wisely depending on the relation which came out after the substitution of the expansion B.8 in the equation $R(\boldsymbol{x}, u; 0) = 0$. The choice of this value affect the importance given to the different terms. Usually, for our purposes it can be choosen $\phi_k = \varepsilon^{k-1}$, as it will be shown in what follow.

Application: Kepler Equaition

Here is presented a very simple application of the concepts explained just above. The equation choosen to be treated is the famous Kepler equation, also known as the time of flight equation for elliptic orbits. It is

$$E - e \sin E = tn \tag{B.9}$$

It is easy to see that if the eccentricity is zero the equation solution is trivial, so one can think to look for a solution $E(t;e)$ for $e \to 0$, which can be useful for near circular orbits since in these cases it is $0 \leq e \ll 1$.

The expansion adopted for the solution is

$$E(t;e) = E_0(t;0) + \phi_2(e) E_2(t) + \phi_3(e) E_3(t) + o(\phi_3(e))$$

and a good choice for the different expansion functions would be $\phi_k = e^{k-1}$, so

$$E(t;e) = E_0(t;0) + eE_2(t) + e^2E_3(t) + o\left(e^2\right)$$

Substituting it into the equation B.9 it gives

$$\left(E_0 + eE_2 + e^2E_3\right) - e\sin\left(E_0 + eE_2 + e^2E_3\right) = tn$$

introducing the McLaurin series for the sine function, considering only the terms up to the second–order and collecting wisely the different quantities one obtains

$$E_0 + e(E_2 - \sin E_0) + e^2(E_3 - E_2\cos E_0) = nt$$

and applying the relation B.3 the different terms can be obtained

$$E_0 = nt \quad E_2 = \sin E_0 \quad E_3 = E_2\cos E_0$$

so the solution found thanks to the expansion is

$$E(t;e) = nt + e\sin(nt) + e^2\sin(nt)\cos(nt)$$

Given that solution the problem is closed evaluating the true anomaly directly from the eccentric anomaly with the relation

$$\cos\nu = \frac{e - \cos E}{e\cos E - 1}$$

In the figure B.3 are plotted the numerical estimation of E (dotted line) and the approximated one (continuos line), and it is quite evident that they are superimposed. In order to highlight the differences between them in the figure B.4 is plotted the error between the two quantities.

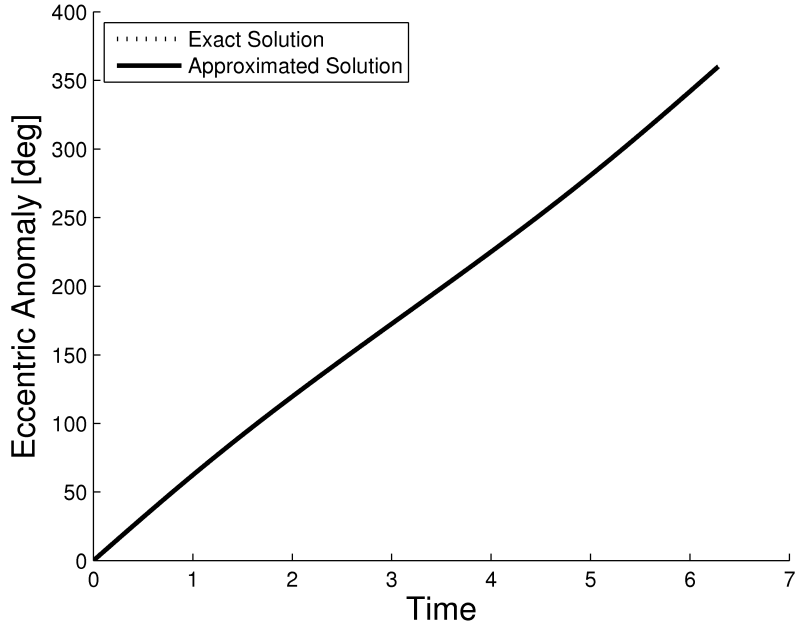

Figure B.3: Comparison between numerical (exact) solution and approximated one

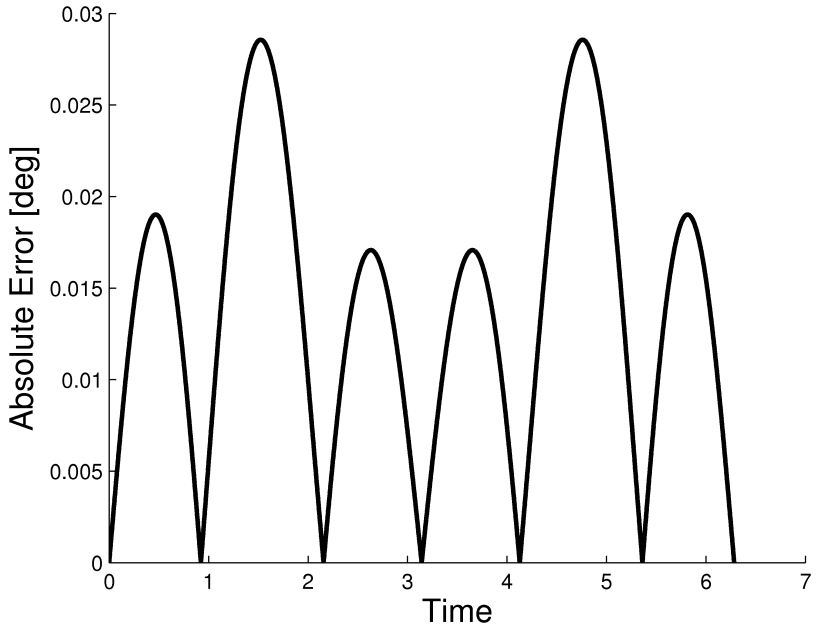

Figure B.4: Error between numerical (exact) solution and approximated one

B.2.2 Regular Expansions for Ordindary Differential Equations

The perturbative expansion can be used also to satisfying the task of finding a function $u(\boldsymbol{x};\varepsilon)$ which is defined as the solutions of ordinary differential equations that involve the parameter ε. There are two kind of situations: the so called regular perturbation problems, where the asymptotic expansion of the solution can be directly derived in a form that remains uniformly valid throughout the domain of interest, and the singular perturbation problems, in which an asymptotic expansion derived by a given limit process often fails somewhere in the domain of interest.

To fix ideas, let L and M be given differential operators. For simplicity assume L to be linear and consider the differential equation

$$L(u) + \varepsilon M(u) = 0 \tag{B.10}$$

in some domain D with initial and/or boundary conditions that do not involve ε. Supposing that $u_0(\boldsymbol{x})$, the solution of

$$L(u_0) = 0$$

satisfying the given initial and/or boundary data, is known. A solution of the perturbed problem B.10 can be assumed in the form

$$u(\boldsymbol{x};\varepsilon) = u_0(\boldsymbol{x}) + \phi_1(\varepsilon)\, u_1(x) + o(\phi_1) \tag{B.11}$$

where $\phi_1(\varepsilon) = o(1)$ as $\varepsilon \to 0$ but is otherwise unknown. Substituting the perturbation expansion in B.10 gives

$$\phi_1(\varepsilon)\, L(u_1) + o(\phi_1) + \varepsilon M(u_0) = 0 \tag{B.12}$$

because $L(u_0 + \phi_1 u_1) = L(u_0) + \phi_1 L(u)$ for a linear operator, and $L(u_0) = 0$. If $\phi_1 = \mathcal{O}(\varepsilon)$, say $\phi = \varepsilon$, u obeys the linear inhomogeneous equation

$$L(u_1) = -M(u0)$$

subject to zero initial and/or boundary data. Because this equation is inhomogeneous, one finds a nontrivial solution $u_1 = (x)$. The other choice, $\varepsilon \ll \phi_1$, is not of interest because it gives $L(u_1) = 0$, and this, along with the vanishing of the initial and/or boundary values, usually implies $u_1 = 0$. The third alternative, $\phi_1 \gg \varepsilon$, leads to an inconsistent condition: it requires that $M(u_0) = 0$, an algebraic relation that is not true in general. Once u_i is calculated using the last equation, the B.11 can be modified to include the

next higherorder term and proceed to derive the equation it obeys. Or one can anticipate the structure of the expansion, say, $\phi_1 = \varepsilon$, $\phi_2 = \varepsilon^2$, ... and solve the sequence of inhomogeneous equations that result from B.10

$$L(u_i) = f_i(\boldsymbol{x}); \; i = 0, 1, 2, ...,$$

where $f_0 = 0$, and each f_i for $i > 0$ is a function of the previously calculated solutions. The following example illustrates these ideas.

Application: Linear Oscillator

As a preliminary application can be choosen the response of a linear Spring-Mass-Damping System, initially at rest, to an impulse I_0 (see fig. B.5). The

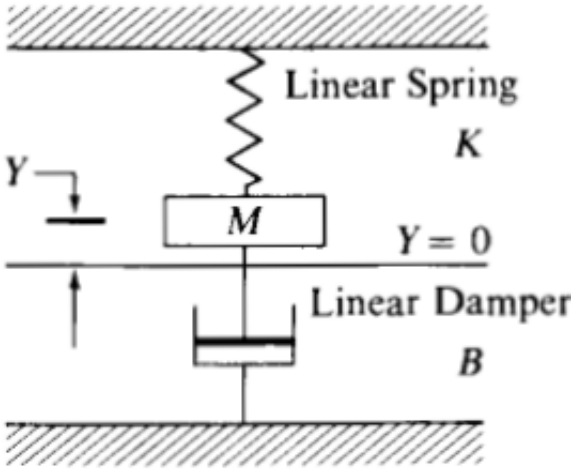

Figure B.5: Spring-Mass-Damping System

equation and initial conditions are

$$M\frac{d^2Y}{dT^2} + B\frac{dY}{dT} + KY = I_0\delta(T) \tag{B.13}$$

$$Y\left(0^-\right) = \frac{dY(0^-)}{dT} = 0 \tag{B.14}$$

where δ is the Dirac delta function. Thi problem can be replaced by an equivalent one by considering an impulse-momentum balance across $T = 0$

or by integrating the previous system from $T = 0^-$ to $T = 0^+$:

$$M\frac{d^2Y}{dT^2} + B\frac{dY}{dT} + KY = 0 \tag{B.15}$$

$$Y\left(0^+\right) = 0 \tag{B.16}$$

$$\frac{dY\left(0^+\right)}{dT} = \frac{I_0}{M} \tag{B.17}$$

The solution defined by this problem is the fundamental solution of this linear equation.

Before proceeding with the perturbation analysis, it is crucial to choose dimensionless variables that are appropriate for the limiting case to be studied. Two such limiting cases are of interest for the linear oscillator: small Damping and small Mass.

Small damping (cumulative perturbation)

If B is small, the motion is expected to be a weakly damped oscillation close to the free simple harmonic oscillation of the system, the solution of B.15 with B = 0. For the introduction of dimensionless coordinates, a suitable time scale is $\sqrt{M/K}$, the reciprocal of the natural frequency of free undamped motion, since this scale remains in the limit $B \to 0$. The length scale A, a measure of the amplitude, can be chosen arbitrarily, and this choice will not affect the resulting dimensionless differential equation since it is linear. Actually, it is convenient to choose A in a form useful to normalizing the initial velocity. Setting

$$t^* = \frac{T}{\sqrt{M/K}} \quad y = \frac{Y}{A}$$

one finds

$$\frac{d^2y}{dt^*} + 2\varepsilon^*\frac{dy}{dt^*} + y = 0 \tag{B.18}$$

where

$$\varepsilon^* = \frac{B}{2\sqrt{MK}}$$

In these variables $y\left(0^+\right) = 0$, $dy(0^+)/dt^* = 1$ if $A = I_0/sqrtMK$. The solution involves the one parameter ε^*, and small damping corresponds to ε^* small. The exact solution is easily found:

$$y\left(t^*, \varepsilon\right) = \frac{e^{-\varepsilon^* t^*}}{\sqrt{1-\varepsilon^*}} \sin\left(\sqrt{1-\varepsilon^{*2}}t^*\right) \tag{B.19}$$

A regular perturbation expansion of this solution, i.e., $\varepsilon^* \to 0$ with t^* fixed and finite, is

$$y = \sin t^* - \varepsilon^* t^* \sin t^* + \mathcal{O}\left(\varepsilon^{*2}\right) + \mathcal{O}\left(\varepsilon^{*2} t^*\right) + \mathcal{O}\left(\varepsilon^{*2} t^{*2}\right) \tag{B.20}$$

This result also follows if one assume the expansion

$$y = g_1\left(t^*\right) + \varepsilon^* g_2\left(t^*\right) + ...$$

and solve the equations that result for g_1 and g_2. The expansion B.20 is uniformly valid to $\mathcal{O}\left(\varepsilon^*\right)$ only if t^* is in the interval $0 \leq t^* \leq T_0 = \mathcal{O}(1)$.

Small mass (singular perturbation)

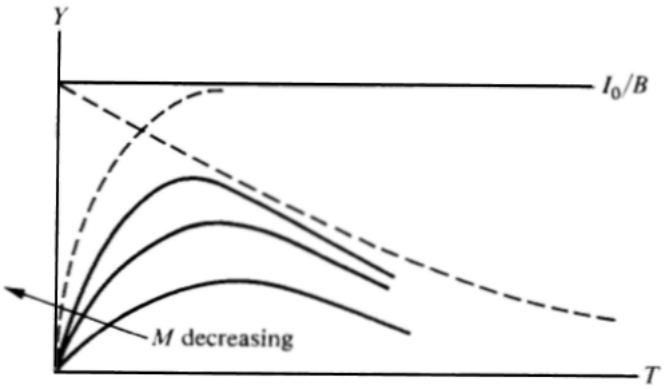

Figure B.6: Solution, varying M

A singular problem is associated with approximations of the starting equations for small values of the mass M. The difficulty near T = 0 arises from the fact that the limit equation with $M = 0$ is first–order, so that the initial conditions cannot both be satisfied. The loss of an initial or boundary condition in a problem leads, in general, to the occurrence of a boundary layer. The general nature of the solution for small values of M is sketched in figure B.6 with each solid curve corresponding to a fixed value of M. After a short time interval, it can be expected that the motion of the system is described by the limit form of the initial equation with $M = 0$:

$$B\frac{dY}{dT} + KY = I_0 \delta\left(T\right) \tag{B.21}$$

The initial condition in velocity is lost, and the effect of the impulse is to introduce a jump in the initial displacement from $Y\left(0^-\right) = 0$ to

$$Y\left(0^+\right) = \frac{I_0}{B}$$

The solution is

$$Y = \frac{I_0}{B} e^{-KT/B} \tag{B.22}$$

The solution decays exponentially after the short initial interval in which the displacement increases infinitely rapidly from 0 to I_0/B. In order to describe the motion during the initial instants, it must be remarked that inertia is certainly dominant at $T = 0$ (impulse-momentum balance). Due to the large initial velocity, damping is immediately important, whereas the restoring force of the spring is not; the spring must be deflected before its influence is felt. Thus, in the initial instants, the starting equation can be approximated by

$$M\frac{d^2Y}{dT^2} + B\frac{dY}{dT} = I_0\delta(t) \tag{B.23}$$

$$Y\left(0^-\right) = 0 \tag{B.24}$$

$$\frac{dY}{dT}\left(0^-\right) = 0 \tag{B.25}$$

with the solution

$$Y(T) = \frac{I_0}{B}\left[1 - e^{-BT/M}\right] \tag{B.26}$$

This solution shows the approach of the deflection in a very short time ($M \to 0$) to the starting value for the decay solution B.22. The curves are shown dashed in figure B.6 and give an overall picture of the motion. With the aim of constructing a suitable asymptotic expansions for expressing these physical ideas and to show how to join these expansions. The method uses expansions valid after a short time (away from the initial point) and expansions valid near the initial point. For the expansion valid away from the initial point, natural variables are those based on a time scale for decay (B/K) and on an amplitude linear in I_0. Let

$$t = \tfrac{K}{B}T \quad y = B\tfrac{Y}{I_0}$$

so the equation becomes

$$\varepsilon\frac{d^2y}{dt} + \frac{dy}{dt} + y = 0 \tag{B.27}$$

where $\varepsilon = MK/B^2$, with initial conditions

$$y(0;\varepsilon) = 0 \quad \tfrac{dy}{dt}(0;\varepsilon) = \tfrac{1}{\varepsilon}$$

The exact solution is

$$y(t;\varepsilon) = \frac{1}{\sqrt{1-4\varepsilon}}\left\{\exp\left[-\left(1-\sqrt{1-4\varepsilon}\right)\frac{t}{2\varepsilon}\right] - \exp\left[-\left(1+\sqrt{1-4\varepsilon}\right)\frac{t}{2\varepsilon}\right]\right\} \tag{B.28}$$

B.3 Multiple Scale Expansions

Various physical problems are characterized by the presence of a small disturbance which, because it is active over a long time, has a non-negligible cumulative effect. For example, the effect of a small damping force over many periods of oscillation is to produce a decay in the amplitude of an oscillator. A more interesting example having the same physical and mathematical features is that of the motion of a satellite around Earth. Here the dominant force is a spherically symmetric gravitational field. If this was the only force acting on the satellite, the motion would be periodic (for sufficiently low energies). The presence of a thin atmosphere, a slightly nonspherical Earth, a small moon, a distant sun, and so on, all produce small but cumulative effects which, after a sufficient number of orbits, drastically alter the nature of the motion. It is the aim of this section to discuss the method of multiple scales, one of the two principal methods for accounting for small cumulative perturbations over a long time. A central feature of the method is the nonexistence, for long times, of a limit process expansion of the type used so extensively in previous examples. As a result, one is led to represent the solution at the outset in the form of a general asymptotic expansion. This is in contrast to the situation encountered in the linear oscillator, where a general asymptotic expansion arose at the last stage of computation when one combined an inner and outer expansion to define the composite solution. Because limit process expansions are not applicable, successive terms in the solution cannot be calculated by the repeated application of limits and, more importantly, rules must be established for the calculation of these terms. Viewed in this light, the method of multiple scales is a generalization of a method proposed by the astronomer Lindstedt for the calculation of periodic solutions.

Application: Weakly nonlinear Oscillator

An elementary example illustrating the basic ideas of the method of multiple scales is that of a linear oscillator with small linear damping. This example was formulated in the previous pages, simplifing the notation

$$\frac{d^2y}{dt^2} + 2\varepsilon\frac{dy}{dt} + y = 0 \tag{B.29}$$

$$y\left(0;\varepsilon\right) = 0 \tag{B.30}$$

$$\frac{dy}{dt}\left(0;\varepsilon\right) = 1 \tag{B.31}$$

where ε is the ratio of the two time scales T_1, T_2:

$$\varepsilon = \frac{T_1}{T_2} = \frac{B}{2\sqrt{KM}}$$

Here T_2 is the damping time

$$T_2 = \frac{2M}{B}$$

which is long if B is small, and $2\pi T_1$ is the period of oscillation for $B = 0$:

$$T_1 = \sqrt{\frac{M}{K}}$$

and T_1 is assumed to be small if compared with T_2.

The physical phenomena occur over these two time scales as can be seen clearly if the exact solution

$$y = \frac{e^{-\varepsilon t}}{\sqrt{1-\varepsilon^2}} \sin \sqrt{1-\varepsilon^2} t \tag{B.32}$$

is written in dimensional variables

$$y = \frac{e^{-\frac{t}{T_2}}}{\sqrt{1-\left(\frac{T_1}{T_2}\right)^2}} \sin \sqrt{1-\left(\frac{T_1}{T_2}\right)^2} \frac{t}{T_1} \tag{B.33}$$

For $\varepsilon \ll 1$ the period of damped oscillations in approximately $2\pi T_1$ and the damping time, which can be defined as the time it takes for the damping to have an $\prime(1)$ effect on the solution, is T_2; an expansion of this solution is

$$y(t;\varepsilon) = \sin t - \varepsilon t \sin t + \mathcal{O}\left(\varepsilon^2\right) + \mathcal{O}\left(\varepsilon^2 t\right) + \mathcal{O}\left(\varepsilon^2 t^2\right) \tag{B.34}$$

This expansion is associated with the limit process $\varepsilon \to 0$, t fixed and is only initially valid $(0 < t < T_0 = \mathcal{O}(1))$ due to the presence of the $\varepsilon t \sin t$ term. In this example, the first mixed-secular term encountered to $\mathcal{O}(e)$ is due to the nonuniform representation for large times of the $e^{-\varepsilon t}$ term in the exact solution. To $\mathcal{O}(\varepsilon^2)$, the expansion of $e^{-\varepsilon t}$ contributes a term proportional to $\varepsilon^2 t^2 sint$; a mixed-secular term proportional to $\varepsilon^2 t \cos t$ will also occur to $\mathcal{O}(\varepsilon^2)$ from the nonuniform representation of the $\sin\sqrt{1-\varepsilon^2}t$ term in the exact solution. It is also evident that mutually contradictory requirements arise if one attempt to represent both εt and $\sin\sqrt{1-\varepsilon^2}t$ uniformly for t in the large interval $I : 0 < t < T(\varepsilon) = \prime(\varepsilon^{-}1)$. In particular, the only uniformly valid representation for $e^{-\varepsilon t}$ in this interval is $e^{-\varepsilon t}$ itself. Therefore,

it is needed the limit process $\varepsilon \to 0$, $\bar{t} = \varepsilon t$ fixed $\neq 0$ in order to represent $e^{-\varepsilon t}$ uniformly in I. However, this limit process does not exist for $\sin\sqrt{1-\varepsilon^2}t = \sin\sqrt{1-\varepsilon^2}\bar{t}/\varepsilon$, as the argument of the sine function tends to infinity as $\varepsilon \to 0$ with $\bar{t} = fixed \neq 0$. Another way of saying this is that the decaying oscillatory function defined by the exact solution does not have an outer limit. On the other hand, as pointed out in [11] Sec. 4.1, it is needed the limit process $\varepsilon \to 0$, $t^+ = (1 - (\varepsilon^2/2) + ...)\, t$ fixed $\neq \infty$, i.e., an expansion in terms of the strained coordinate t^+, in order to uniformly represent $\sin\sqrt{1-\varepsilon^2}t$ over I. In this case $e^{-\varepsilon t}$ is expressed as $e^{-\varepsilon t^+\left(1+\varepsilon^2/2+...\right)}$ and leads to essentially the same nonuniformity in I as the initially valid expansion B.34.

General asymptotic expansion, two scale expansion

Any asymptotic expansion of the previous exact solution must simultaneously depict both the decaying and oscillatory behaviors of the solution in order to be uniformly valid in I. It is clear that a limit process expansion will not do so it is needed to look for a general asymptotic expansion where each term depends on t and ε. In fact, if one avoid expanding $e^{-\varepsilon t}$ and simply develop the argument $\sqrt{1-\varepsilon^2}t$ without further expanding the sine function, it can be found the following general asymptotic expansion for the exact solution in a form that is uniformly valid in I:

$$y(t;\varepsilon) = \sum_{n=0}^{N} \left(\alpha_n e^{-\varepsilon t} \sin\Omega_n(\varepsilon)\, t\right) \varepsilon^n + \mathcal{O}\left(\varepsilon^{N+1}\right) \tag{B.35}$$

Here α_n follows from the expansion of the factor $(1-\varepsilon^2)^{-\frac{1}{2}}$ in the exact solution, whereas Ω_n results from expanding the frequency $(1-\varepsilon^2)^{\frac{1}{2}}$

$$\alpha_{2n} = \frac{1}{2^n n!} \prod_{k=1}^{n+1} |2k-3|; \quad n = 0, 1, ..., \tag{B.36}$$

$$\alpha_{2n+1} = 0; \quad n = 0, 1, 2, ..., \tag{B.37}$$

$$\Omega_0(\varepsilon) = 1 \tag{B.38}$$

$$\Omega_{2n-1}(\varepsilon) = 1 - \sum_{j=1}^{n} \frac{1}{2^j j!} \prod_{k=1}^{j} |2k-3|\, \varepsilon^{2j}; \quad n = 1, 2, ..., \tag{B.39}$$

$$\Omega_{2n} = \Omega_{2n-1}; \quad n = 1, 2, ..., \tag{B.40}$$

$$\tag{B.41}$$

In particular, the uniformly valid approximations to order $1, \varepsilon, \varepsilon^2, \varepsilon^3, ...$, denoted respectively as $y^{(0)}$, $y^{(1)}$, $y^{(2)}$, $y^{(3)}$, ..., are

$$y^{(0)} = e^{-\varepsilon t} \sin t \tag{B.42}$$

$$y^{(1)} = e^{-\varepsilon t} \sin \left(1 - \frac{\varepsilon^2}{2}\right) t \tag{B.43}$$

$$y^{(2)} = \left(1 + \frac{\varepsilon^2}{2}\right) e^{-\varepsilon t} \sin \left(1 - \frac{\varepsilon^2}{2}\right) t \tag{B.44}$$

$$y^{(3)} = \left(1 + \frac{\varepsilon^2}{2}\right) e^{-\varepsilon t} \sin \left(1 - \frac{\varepsilon^2}{2} - \frac{3\varepsilon^4}{8}\right) t \tag{B.45}$$

$$\tag{B.46}$$

As pointed out in Problem 6 of Sec. 4.1 [11], it is needed to account for the frequency to $\mathcal{O}\left(\varepsilon^{N+1}\right)$ in order to have uniform validity in I to $\mathcal{O}\left(\varepsilon^{N}\right)$. In this example, the frequency expansion proceeds in even powers of ε.

It can be noted also that the general asymptotic expansion B.35 to any $\mathcal{O}\left(\varepsilon^{N}\right)$ may be expressed uniquely in the form of a series of functions of tine fast scale $t_N^+ = \Omega_N(\varepsilon)$ and the slow scale $\bar{t} = \varepsilon t$. In fact, the exact expression is itself a unique function of t_∞^+ and $\bar{t}$ where $t_\infty^+ = \Omega_\infty(\varepsilon)\, t = \sqrt{1-\varepsilon^2}t$.

So it is

$$y(t;\varepsilon) = F\left(t_\infty^+, \bar{t}; \varepsilon\right)$$

where

$$F\left(t_\infty^+, \bar{t}; \varepsilon\right) = \frac{e^{-\bar{t}}}{\sqrt{1-\varepsilon^2}} \sin t_\infty^+$$

and for a given integer N, F as a unique two-scale expansion correct to $\mathcal{O}\left(\varepsilon^{N}\right)$ uniformly in I of the form

$$F = \sum_{n=0}^{N} F_n\left(t_N^+, \bar{t}\right) \varepsilon^n + \mathcal{O}\left(\varepsilon^{N+1}\right) \tag{B.47}$$

where for N even is

$$F_0 = e^{-\bar{t}} \sin t_N^+ \tag{B.48}$$

$$F_1 = 0 \tag{B.49}$$

$$F_2 = \frac{e^{-\bar{t}}}{2} \sin t_N^+ \tag{B.50}$$

$$F_3 = 0 \tag{B.51}$$

$$\vdots \tag{B.52}$$

$$F_N = \alpha_N e^{-\bar{t}} \sin t_N^+ \tag{B.53}$$

$$t_N^+ = \Omega_N(\varepsilon)\, t \tag{B.54}$$

Two-scale expansion of the differential equation

Guided by the above results, lets try to reconstruct the expansion B.47 from the governing differential equation and initial conditions without direct knowledge of the exact solution, this idea will later be applied to nonlinear problems where the exact solution is not available. The fundamental assumption is that solutions have a general asymptotic expansion that is uniformly valid in the interval I: $0 < t < T(\varepsilon) = \mathcal{O}(\varepsilon^{-1})$, and each term in this general asymptotic expansion can be uniquely expressed as a function of t^+, t and a power of ε as in B.47, this assumption allows to calculate uniformly valid expansions in a wide variety of problems. Examining the implications of this assumption: suppose to encounter a mixed-secular term of the form $\varepsilon t_2^+ \sin t_2^+$ in the $\mathcal{O}(\varepsilon)$ contribution to the expansion, where $t_2^+ = (l + \varepsilon\omega_1 + \varepsilon^2\omega)\, t$. Such a term is not consistent with a uniquely defined two scale expansion because this term can be expressed as follows:

$$\varepsilon t_2^+ \sin t_2^+ = \bar{t} \sin t_2^+ + \varepsilon\omega_1 \bar{t} \sin t_2^+ + \varepsilon^2 \omega_2 \bar{t} \sin t_2^+ \tag{B.55}$$

and redistribute its contribution to various orders. A term need not become unbounded as $t \to \infty$ to be inconsistent. For example, the exponentially decaying term $\varepsilon\bar{t}e^{-\bar{t}} \sin t_2^+$ is also unacceptable because one can always relabel $\varepsilon\bar{t}$ in terms of t_2^+ and $\bar{t}$ and change its nominal order, i.e.,

$$\varepsilon\bar{t} = \varepsilon^2 t_2^+ - \varepsilon^2\omega_1\bar{t} - \varepsilon^3\omega_2\bar{t} + ...$$

We note, however, that it is not possible to uniquely allocate an $\mathcal{O}(\varepsilon)$ contribution in the frequency in terms of a t^+ and $\bar{t}$ contribution. To see this,

consider a term of the form

$$g = \sin\left(1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3\right) t$$

for given nonzero constants $\omega_1, \omega_2, \omega_3$. With t_3^+ defined as

$$\left(1 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3\right) t$$

such a term would simply be denoted $g = \sin t_3^+$. However, if another fast scale it is choosen, say, $t_3^* = (1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3)\, t$ instead of t_3^+, then it is found $g = \cos\omega_1\bar{t}\sin t_3^* + \sin\omega_1 t\cos t_3^*$. In fact, there are an infinite number of possible choices of fast scale corresponding to different choices of the $\mathcal{O}(\varepsilon)$ term in the expansion of the frequency Ω; each of these choices results in a different two-scale representation for g. To avoid this ambiguity, we will henceforth set $\omega_1 = 0$ in the definition of t^+ and account for it via the $\bar{t}$-dependence of the solution. Looking for a two-scale expansion of the solution of the initial system in the form

$$y(t;\varepsilon) = F\left(t^+, \bar{t}; \varepsilon\right) = \sum_{n=0}^{N} F_n\left(t^+\bar{t}\right)\varepsilon^n + \mathcal{O}\left(\varepsilon^{\mathcal{N}+\infty}\right) \tag{B.56}$$

Here

$$t^+ = \left(1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3 + ...\right) t \tag{B.57}$$

where $\omega_2, \omega_3, ...$ are unknown constants, and $\bar{t} = \varepsilon t$. The chain rule gives

$$\frac{dy}{dt} = \left(1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3 + ...\right)\frac{\partial F}{\partial t^+} + \varepsilon\frac{\partial F}{\partial \bar{t}} \tag{B.58}$$

$$\frac{d^2y}{dt^2} = \left(1 + \varepsilon^2\omega_2 + \varepsilon^3\omega_3 + ...\right)^2\frac{\partial^2 F}{\partial t^{+2}} + \tag{B.59}$$

$$+ \quad 2\varepsilon\left(1 + \varepsilon^2\omega_2 + \varepsilon^2\omega_3 + ...\right)\frac{\partial^2 F}{\partial t^+\partial\bar{t}} + \varepsilon^2\frac{\partial^2 F}{\partial\bar{t}^2} \tag{B.60}$$

and using the expansion for F it is found to $\mathcal{O}(\varepsilon^2)$

$$\frac{dy}{dt} = \frac{\partial F_0}{\partial t^+} + \varepsilon\left(\frac{\partial F_1}{\partial t^+} + \frac{\partial F_0}{\partial\bar{t}}\right) + \varepsilon^2\left(\frac{\partial F_2}{\partial t^+} + \frac{\partial F_1}{\partial\bar{t}} + \omega_2\frac{\partial F_0}{\partial t^+}\right) \tag{B.61}$$

$$\frac{d^2y}{dt^2} = \frac{\partial^2 F_0}{\partial t^{+2} + \varepsilon\left(\frac{\partial^2 F_1}{\partial t^{+2}} + 2\frac{\partial^2 F_0}{\partial^+\partial\bar{t}}\right)} + \tag{B.62}$$

$$+ \quad \varepsilon^2\left(\frac{\partial^2 F_2}{\partial t^{+2}} + 2\frac{\partial^2 F_1}{\partial t^+\partial\bar{t}} + 2\omega_2\frac{\partial^2 F_0}{\partial t^{+2}} + \frac{\partial^2 F_0}{\partial\bar{t}^2}\right) \tag{B.63}$$

Thus, the sequence of equations that results from the initial equation is

$$L(F_0) \equiv \frac{\partial^2 F_0}{\partial t^{+2}} + F_0 = 0 \tag{B.64}$$

$$L(F_1) \equiv -2\frac{\partial^2 F_0}{\partial t^+ \partial \bar{t}} - 2\frac{\partial F_0}{\partial t^+} \tag{B.65}$$

$$L(F_2) \equiv -2\omega_2 \frac{\partial^2 F_0}{\partial t^{+2}} - \frac{\partial^2 F_0}{\partial \bar{t}^2} - 2\frac{\partial^2 F_1}{\partial t^+ \partial \bar{t}} - 2\frac{\partial F_0}{\partial \bar{t}} - 2\frac{\partial F_1}{\partial t^+} \tag{B.66}$$

The first of these is the equation for the free oscillations, while the remainder have the appearance of forced linear oscillations. However, since $F_0 = F_0(t^+, \bar{t})$, the free linear oscillations that are the solutions to B.64 have the possibility of being slowly modulated. Thus, it can be written

$$F_0\left(t^+, \bar{t}\right) = A_0(\bar{t}) \cos t^+ B_0(\bar{t}) \sin t^+ \tag{B.67}$$

imposing the initial conditions and using the derivatives rules previuos explained one has

$$\begin{array}{ll} F_0(0,0) = 0 & \frac{\partial F_0}{\partial t^+}(0,0) = 1 \\ F_1(0,0) = 0 & \frac{\partial F_1}{\partial t^+} = -\frac{\partial F_0}{\partial \bar{t}}(0,0) \\ F_2(0,0) = 0 & \frac{\partial F_2}{\partial t^+}(0,0) = -\frac{\partial F_1}{\partial \bar{t}}(0,0) - \omega_2 \frac{\partial F_0}{\partial t^+}(0,0) \end{array}$$

and these equation leads to the following conditions

$$A_0(0) = 0, \quad B_0(0) = 1 \tag{B.68}$$

Nothing more can be found out about $A_0(\bar{t})$ and $B_0(\bar{t})$ without considering F_1. This is directly analogous to the situation encountered in Sec. 4.1 of [11] for the method of strained coordinates. Substituting for F_0 into the right-hand side of B.65 gives

$$L(F_1) = 2\left[\frac{dA_0}{d\bar{t}} + A_0\right] \sin t^+ - 2\left[\frac{dB_0}{d\bar{t}} + B0\right] \cos t^+ \tag{B.69}$$

The bracketed terms on the right-hand side of this equation are functions of $\bar{t}$ only. Therefore, the particular solutions corresponding to these terms would be functions of $\bar{t}$ multiplied by the mixed-secular terms $t^+ \sin t^+$ or $t^+ \cos t^+$. Such terms cannot be permitted to occur in the solution because, as discussed earlier, they are inconsistent with a unique F_1.

Therefore, it is needed to eliminate all homogeneous solutions of $L(F_1) = 0$ from the right-hand side of the last equation, and this gives the two first-order ordinary differential equations for A_0 and B_0:

$$\frac{dA_0}{d\bar{t}} + A_0 = 0 \tag{B.70}$$

$$\frac{dB_0}{d\bar{t}} + B_0 = 0 \tag{B.71}$$

Taking account of the initial conditions, it is found that

$$A_0(\bar{t}) = 0, \quad B_0(\bar{t}) = e^{-\bar{t}} \tag{B.72}$$

The uniformly expansion to $\mathcal{O}(1)$ is

$$y(t;\varepsilon) = e^{-\bar{t}} \sin t^+ \mathcal{O}(\varepsilon) \tag{B.73}$$

where $t^+ = (1 + \mathcal{O}(\varepsilon^2))\, t$ and this agrees with the exact result (see B.42). Thus far, it has been determined the first two terms in the expansion

$$F_0\left(t^+, \bar{t}\right) = e^{-\bar{t}} \sin t^+ \tag{B.74}$$

$$F_1\left(t^+, \bar{t}\right) = A_1(\bar{t}) \cos t^+ B_1(\bar{t}) \sin t^+ \tag{B.75}$$

and finally F_2 becomes

$$L(F_2) = \left[2\left(\frac{dA_1}{d\bar{t}} + A_1\right) + (2\omega_2 + 1)\, e^{-\bar{t}}\right] \sin t^+ - 2\left[\frac{dB_1}{d\bar{t}} + B_1\right] \cos t^+ \tag{B.76}$$

Now, $A_1(\bar{t})$, $B_1(\bar{t})$, and the frequency shift ω_2 are to be found from similar considerations applied to this last equation.

First, repeating the argument that homogeneous solutions of $L(F_2) = 0$ cannot be permitted, it must be setted the bracketed terms in the last equation equal to zero. Solving the resulting equations for A_1 and B_1 subject to the initial conditions $A_1(0) = B_1(0) = 0$, it is found

$$A_1(\bar{t}) = -\left(\omega_2 + \frac{1}{2}\right) \bar{t} e^{-\bar{t}} \tag{B.77}$$

$$B_1(\bar{t}) = 0 \tag{B.78}$$

This means that εF_1 would contain a term proportional to $\varepsilon \bar{t} e^{-\bar{t}} \cos t^+$. Again, such a term cannot be consistent because, as pointed out earlier, it can also be written as $\varepsilon^2 e^{-\bar{t}} t^+ \cos t^+ \mathcal{O}(\varepsilon^3)$ and shift to $\mathcal{O}(\varepsilon^2)$ in the expansion. One

can also have required that $|F_2/F_1|$ be bounded for large $\bar{t}$ to disallow such a term. Therefore, it must be setted

$$\omega_2 = -\frac{1}{2}$$

and the following uniformly valid result in I to $\mathcal{O}(\varepsilon)$ is found

$$y(t;\varepsilon) = e^{-\bar{t}} \sin\left[1 - \frac{\varepsilon^2}{2} + \mathcal{O}\left(\varepsilon^3\right)\right] t + \mathcal{O}\left(\varepsilon^2\right) \tag{B.79}$$

All the necessary reasoning has now been explained to carry out the solution to any order and, in fact, to solve a wide variety of weakly nonlinear problems of the form

$$\frac{d^2y}{dt^2} + y + \varepsilon f\left(y, \frac{dy}{dt}\right) = 0 \tag{B.80}$$

Appendix C

Fundamentals of Optimization Methods

In this chapter a brief general overview on optimization methods is given, first of all there is a classification of optimization tecniques and then the description of each category, but only for the direct methods are reported some details, since this category is the one used for apply the perturbative expansions previously derived.

C.1 Classification

Optimization algorithms are usually divided in three groups

- Indirect Optimization Methods;
- Evolutionistic Methods;
- Direct Optimization Methods.

Each category has its peculiarity which fit some kind of problems better than others, their description is given in the next sections.

C.2 Indirect Optimization Methods

These methods can face continous problems, so they are not affected by the problem of the discretization. First of all must be defined the following quantities:

- t indipendent variable;

- $\boldsymbol{x}$ state variables (n-component vector);
- $\boldsymbol{u}$ control variables (m-component vector);
- $\dot{\boldsymbol{x}} = dx/dt = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t)$ state equations (n-component vector of differential equations);
- $\boldsymbol{\psi}(\boldsymbol{x}_0, \boldsymbol{x}_f, t_0, t_f) = 0$ boundary conditions (q-component vector of algebraic equations, $q \leq n + 2$)

and the issue is to find the so called extremal path (i.e. the trajectory) $\boldsymbol{x}(t)$ and the corresponding optimal control law $\boldsymbol{u}(t)$ satisfying the state equations and the contraints to maximize (or minimize, it is the same[1])the functional, i.e. performance index,

$$J = \phi(\boldsymbol{x}_0, \boldsymbol{x}_f, t_0, t_f) + \int_{t_0}^{t_f} \Phi(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)\, dt \tag{C.1}$$

What the indirect methds do, as their name suggest, is to obtain the condition for the maximum of the performance index without evaluating it directly. In order to work on a functional which automatically takes into account the constraints it is useful/needed to build the augmented performance index

$$J^* = J + \boldsymbol{\mu}^T \boldsymbol{\psi} + \int_{t_0}^{t_f} \left[\boldsymbol{\Phi} + \boldsymbol{\lambda}^T (\boldsymbol{f} - \dot{\boldsymbol{x}})\right] dt \tag{C.2}$$

where $\boldsymbol{\lambda}$ are the adjoint variables and $\boldsymbol{\mu}$ are the adjoint constants; in this way, when the contraints are satisfied

$$J^* = J$$

In order to obtain the optimality conditions one has to analize the first variation of the augmented index, so

$$\begin{aligned} dJ^* &= \left(\frac{\partial \phi}{\partial t_f} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial t_f} + H_f\right) dt_f + \\ &+ \left(\frac{\partial \phi}{\partial t_0} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial t_0} - H_0\right) dt_0 + \\ &+ \left(-\boldsymbol{\lambda}_f^T + \frac{\partial \phi}{\partial \boldsymbol{x}_f} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{x}_f}\right) d\boldsymbol{x}_f + \\ &+ \left(\boldsymbol{\lambda}_0^T + \frac{\partial \phi}{\partial \boldsymbol{x}_0} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{x}_0}\right) d\boldsymbol{x}_0 \\ &+ \int_{t_0}^{t_f} \left[\left(\frac{\partial H}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\lambda}}^T\right) \delta \boldsymbol{x} + \frac{\partial H}{\delta \boldsymbol{u}} \delta \boldsymbol{u}\right] dt \end{aligned}$$

[1]Minimize the time t can be seen as maximizing $-t$.

in the maximum point the first variation must be equal to zero and isolating each term one obtains: two boundary conditions for optimality ($2n$ algebraic equations at initial and final point)

$$-\boldsymbol{\lambda}_f^T + \frac{\partial \phi}{\partial \boldsymbol{x}_f} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{x}_f} = 0; \quad \boldsymbol{\lambda}_0^T + \frac{\partial \phi}{\partial \boldsymbol{x}_0} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{x}_0} = 0 \tag{C.3}$$

two transveraslity conditions (2 algebraic equations at initial an final time)

$$\frac{\partial \phi}{\partial t_f} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial t_f} + H_f = 0; \quad \frac{\partial \phi}{\partial t_0} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{\psi}}{\partial t_0} - H_0 = 0$$

one set of differential equations for the adjoint variables, Eulero-Lagrange equations (n differential equations)

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \boldsymbol{x}}^T$$

and one set of equations for the optimal control (m algebraic equations for the control variables)

$$\frac{\partial H}{\partial \boldsymbol{u}}^T = 0$$

So the problem is completely defined and all the variable trends in time can be determined solving numerically the differential equations. Besides this reasoning there are also other kind of evaluations, like the controllability condition and the second variation evalutation, but here they are not taken into account. The main problems for these methods are the possibility of finding suboptimal solutions instead of globally optimal ones, and the dependence on the initial guess for the convergence.

C.3 Evolutionistic Methods

These methods can deal only with discrete systems and also a small number of variables (maximum twenty), in fact they are applied in astrodynamics only with impulsive thrust trajectory because of the need of discretization that make impossible their use in low-thrust problems.

There are a lot of them, here after are listed the principal ones.

Genetic Algorithms

This category emulate the human evolution, a set of individuals (the set of possible solutions) are randomly initialized then their own performance index

is evaluated and the natural selection is performed taking this quantity as discriminant between who is suitable for the given ambient (the optimization problem) and who is not; the selection method can be done in many different ways, for example two ways are the tournament and the roulette: the first one consist in compare the solutions in couples two times each, the winner (who has the best index) in putted in the parents group and the looser is not; in the roulette method the probability clove is given proportional to the performance index and then the extraction is take randomly. After this selection the best individuals are used to generate a new younger genration of solutions, by means of the crossover function, a particular function which decide randomly how near or far the sons would be respect to the parents; then the algorithm is repeated until the value of a certain parameter (number of iterations, increment of performance index, etc.) is matched.

There are some particular techniques which can be used for avoid the local minimun, and they are inspired by the nature too, so sometimes the worse individual are allowed to reproduce, sometimes there is a mass extinction of individuals and only few of the best survive, other times are mutations are allowed, etc.

Particle Swarm Optimization

This algorithm simulate fishes or birds looking for food, a randomly initialized set of solution is created and this time, instead of being replaced by other solutions, the individuals move in the domain of the problem. Their motion is assumed to be similar to the one adopted by a flock of animals when looking for food, in particular the acceleration of the individuals is given by two terms

$$\begin{aligned} v &= v + c_1 k_1 \left(x_{P,best} - x\right) + c_2 k_2 \left(x_{G,best} - x\right) \\ y &= x + v \end{aligned}$$

the first is the personal-cognitive one (which points towards the area where the best value of performance index has been found in the previous steps) and the second is the global-social one (which points torwards the area where the best value of performance index has been found by the whole group); also in this case there are some strategies for impoving the method, for example for the choice of the coefficients of the acceleration and in the choice of associating an intertia to the old velocity.

Differential Evolution

This method simulate the evolution too, a randomly generated set of solutions is created but in this case the process is more simple and the new population is created adding to one individual (it can be the best one or not) a component proportional (by means of some constants which can be used as degrees of freedom) to the difference between other two (or more than two) individuals:

$$y = x_1 + C\left(x_2 - x_3\right) + F\left(x_4 - x_5\right)$$

and the choice of x_i characterize the method. In this way when the differences are big (at the beginning of the search) the domain is explored widely, while when the solutions are near the real optimum the search is performed only in that restricted area.

There are a lot of other method, like for example the ant colony optimization which simulate the motion of the ants (very useful in vehicle routing problems), the simulated annealing which simulate the behaviour of metallic crystals that hare first heated and then cooled, the invasive weed optimization which emulate the weed expansion when looking for a suitable ambient, etc.

Since the initialization is done randomly and the search method is not based on standard optimality conditions these methods should not fall in sub-optimal solutions, but there is not any certainty of convergence, in fact often the method is evaluated looking to the percentage of times it converges (in test case for which the optimal solution in known) on the total of the runs.

C.4 Direct Optimization Methods

In these methods the perturbative expansions will be applied so the description of them would be a little bit more detailed. They need to discretize the trajectory, so this is an optimization of discrete systems: performance index $\phi\left(x\right)$ must be maximized with the contraints $\boldsymbol{c}\left(\boldsymbol{x}\right) \geq 0$, where

- $\boldsymbol{x}$ is a n-component vector of variables, $\boldsymbol{x}^T = [x_1, x_2, ..., x_n]$;
- $\boldsymbol{c}$ is a m-component vector of constraints, $\boldsymbol{c}^T = [c_1, c_2, ..., c_m]$, and m maybe larger than n;

if both the performance index and the constraints are linear this problem is called Linear Programming, and a tipical solving method si the simplex one, but since in space design the relations are always nonlinear here it is not analized.

When ϕ and $\boldsymbol{c}$ are nonlinear the maximization problem become the so called Nonlinear Programming, usually it is faced with various gradient methods which differ for the choice of the step and some other features, but here we refer to the so called Sequential Quadratic Programming - SQP.

Sequential Quadratic Programming

It is based on two approximations: one second–order approximation for the performance index and a first–order one for the constraints and the maximization problem solution is found by the solution of a succession of quadratic programming problems starting from a tentative solution $\boldsymbol{x}$. Assuming those approximations one can write

$$\phi(\boldsymbol{x}+\Delta\boldsymbol{x}) = \phi(\boldsymbol{x}) + \boldsymbol{g}^T\Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^T[\boldsymbol{H}]\Delta\boldsymbol{x} \tag{C.4}$$

$$\boldsymbol{c}(\boldsymbol{x}+\Delta\boldsymbol{x}) = \boldsymbol{c}(\boldsymbol{x}) + [\boldsymbol{G}]\Delta\boldsymbol{x} \tag{C.5}$$

where the linear approximation for the gradient is adopted and

$$\boldsymbol{g}(\boldsymbol{x}+\Delta\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}) + [\boldsymbol{H}]\Delta\boldsymbol{x}$$

where $[\boldsymbol{H}]$ is the Hessian Matrix.

In order to treat the constraints the active set method can be used, it implies to consider the augmented performance index and to consider some constraints as active ($\boldsymbol{c}_a$) and some inactive ($\boldsymbol{c}_{na}$), so

$$\phi^* = \phi(\boldsymbol{x}) + \boldsymbol{\lambda}_a^T\boldsymbol{c}_a$$

and considering an increment one has

$$\phi^*(\boldsymbol{x}+\Delta\boldsymbol{x}) = \phi^*(\boldsymbol{x}) + \left(\boldsymbol{g}^T + \boldsymbol{\lambda}_a^T[\boldsymbol{G}_a]\right)\Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^T[\boldsymbol{H}]\Delta\boldsymbol{x}$$

where the augmented Hessian is

$$[\boldsymbol{H}^*] = [\boldsymbol{H}] + \partial\left[\partial\left(\boldsymbol{\lambda}_a^T\boldsymbol{c}_a/\partial\boldsymbol{x}\right)^T\right]/\partial\boldsymbol{x}$$

and the conditions for optimality (for the given active set) are

$$\begin{aligned} \boldsymbol{g} + [G_a]^T\boldsymbol{\lambda}_a &= 0 \\ \boldsymbol{c}_a &= 0 \end{aligned}$$

and the addition of

$$\Delta\boldsymbol{x}^T[\boldsymbol{H}]\Delta\boldsymbol{x} < 0$$

for any variation $\Delta \boldsymbol{x}$ makes this set of condition sufficient for a maximum.

In general these conditions are not satisfied in a generic point so looking for a new point $\boldsymbol{x} + \Delta \boldsymbol{x}$, $\boldsymbol{\lambda}_a + \Delta \boldsymbol{\lambda}_a$ one can write

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{x}) + [\boldsymbol{H}^*] \Delta \boldsymbol{x} + [\boldsymbol{G}_a]^T (\boldsymbol{\lambda}_a + \Delta \boldsymbol{\lambda}_a) &= 0 \\ \boldsymbol{c}_a(\boldsymbol{x}) + [\boldsymbol{G}_a] \Delta \boldsymbol{x} &= 0 \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{x} + \Delta \boldsymbol{x}) &= \boldsymbol{g}(\boldsymbol{x}) + [\boldsymbol{H}^*] \Delta \boldsymbol{x} \\ c_a(\boldsymbol{x} + \Delta \boldsymbol{x}) &= \boldsymbol{c}_a(\boldsymbol{x}) + [\boldsymbol{G}_a] \Delta \boldsymbol{x} \end{aligned}$$

and the previous equations can be rewritten in the most known form of the so called Karush - Kuhn - Tucker equations

$$\begin{bmatrix} [\boldsymbol{H}^*] & [\boldsymbol{G}_a]^T \\ [\boldsymbol{G}_a] & 0 \end{bmatrix} \begin{pmatrix} \Delta \boldsymbol{x} \\ \boldsymbol{\lambda}_a + \Delta \boldsymbol{\lambda}_a \end{pmatrix} = \begin{pmatrix} -\boldsymbol{g} \\ -\boldsymbol{c}_a \end{pmatrix}$$

and the search for the new point become an iterative procedure.

In the active set method one must not violate the inactive constraints so the step in the gradient direction towards the maximum point can be limited by a factor $\alpha \in [0, 1]$ where

$$\alpha = min \left(\frac{-c_{na,i}}{\frac{\partial c_{na,i}}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}}, 1 \right)$$

when $\alpha < 0$ it is determined by $c_{na,i} = 0$ so the costrain must be added to the active set and when $\lambda_{a,i} < 0$ the corresponding constrain must be removed from the active set.

Until now the method can be regarded as a simple Newton's method, but there are some strategies for improving the algorithm, these are the main differences respect to the indirect methods; in particular this task is accomplished looking at the performance index in order to see if it is growing or not. A merit function is defined and in order to take into account also the constraints it can be defined as

$$M = \phi + \boldsymbol{\lambda}_a^T \boldsymbol{c}_a - \frac{1}{2} \boldsymbol{c}_a^T [\boldsymbol{\rho}] \boldsymbol{c}_a$$

and it is the augmented Lagrangian merit function, with ρ diagonal matrix with non negative terms: once $\Delta \boldsymbol{x}$ and $\Delta \boldsymbol{\lambda}$ are calculated the merit function is only a function of α so it is possible to chose the value of this parameter for which the merit function is maximum; this method is the so called linear

search. There is also the trust region method which associate at the approximation of the index and the constraints a region of validity and doing this another equation cames out

$$\frac{1}{2}\Delta \boldsymbol{x}^T \Delta \boldsymbol{c} \leq \delta^2$$

so the KKT equations has to be modified introducing instead of the augmented Hessian matrix the term $[\boldsymbol{H}^*] - \tau\,[\boldsymbol{I}]$.

For computational reasons linked to the calculus of the matrix and their factorization, can be useful to introduce the slack variables, in this way one can transform inequality constraints in equality ones and force the inequality on the slack variables.

As said at the beginning of the section direct optimization can only deal with discrete systems, so the trajectory must be splitted into finite arcs, for example if the time is the indipendent variable it must be splitted in small intervals $t_{i+1} - t_i = h_i$ where h_i is the integration step and an algebraic approximation must be introduced $\boldsymbol{y}_{i+1} = \boldsymbol{y}_i + \int_{t_i}^{t_{i+1}} \boldsymbol{f}\,(\boldsymbol{y}, t)\,dt$. This relation can be translated in terms of an integration scheme (Runge Kutta, Trapezoidal, Hermite Simpson, etc.). The usual way of solving these problems is tho use the Direct Transcription, it implies the use of a Multiple Shooting technique: the variables are the state and the controls at each node of any subinterval previously defined, supplementary equations are the matching conditions between the final point of one interval (which depends on the initial values on the node of that small subelement and the corresponding states and controls equations) and the initial one on the next element (see next figure for clearness). In order to evaluate the various matrix (gradient,

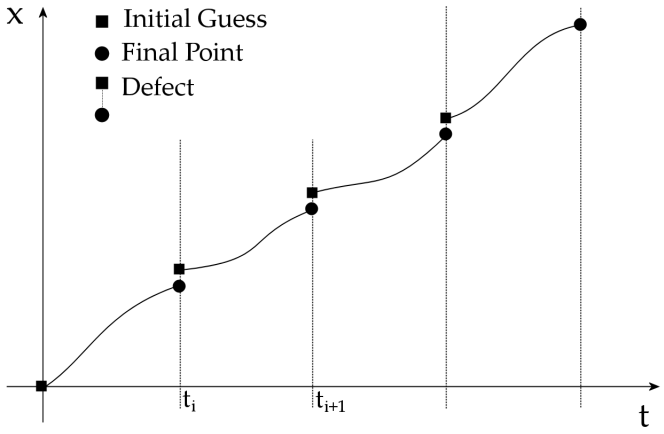

Figure C.1: How multiple shooting works

hessian, etc.) a perturbation on the initial parameters is introduced on the first integration step and then it is applied the recursive procedure.

Bibliography

[1] R.R. Bate, D.D. Mueller, e J.E. White, *Fundamentals of Astrodynamics*, Dover Pubblications Inc., New York, 1971.

[2] R.H. Battin, *Introduction to Mathematics and Methods in Astrodynamics*, AIAA Education Series, Reston (VA), 1999.

[3] C.L. Bottasso, A. Ragazzi, "Finite Element and Runge-Kutta Methods for Boundary-Value and Optimal Control Problems", *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 4, Engineering Notes, pp. 749-751.

[4] A.E. Bryson Jr, Y.C. Ho, *Applied Optimal Control*, Taylor & Francis Group, New York, 1975.

[5] L. Casalino, S. Campagnola, "Trajectories towards near-earth-objects using solar electric propulsion", *AAS/AIAA Astrodynamics Conference*, AAS Paper 99-339.

[6] J.W. Cornelisse, *Rocket Propulsion and Spaceflight Dynamics*, Pitman Publishing, London, 1979.

[7] S.S. Fernandes, F.C. Carvalho, "A First-Order Analytical Theory for Optimal Low-Thrust Limited-Power Transfers between Arbitrary Elliptical Coplanar Orbits", *Hindawi Publishing Corporation Mathematical Problems in Engineering*, Vol. 2008, Article ID 525930.

[8] D. Izzo, "Lambert's problem for exponential sinusoids", *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 5, 2006, pp. 1242-1245.

[9] D. Izzo, R. Bevilacqua, C. Valente "Internal Mesh Optimisation and Runge-Kutta Collocation in a Direct Transcription Method Applied to Interplanetary Missions", *55th International Astronautical Congress*, Paper IAC-04-A.6.04.

[10] R.S. Johnson, *Singular Perturbation Theory*, Springer, Boston, 2005.

[11] J. Kevorkian, J.D. Cole, *Multiple Scale and Singular Perturbation Methods*, Springer, New York, 1996.

[12] J. Kevorkian, M. Eckstein, Y. Shi, "Satellite Motion for All Inclinations Around an Oblate Planet", *The Theory of Orbits in the Solar System and in Stellar Systems*, IAU Symposium No. 25, (Academic Press Inc., New York, 1966), p. 291.

[13] J. Kevorkian, M. Eckstein, Y. Shi, "Satellite Motion for Arbitrary Eccentricity and Inclination around the Smaller Primary in the Restricted Three-Body Problem.", *The Astronomical Journal*, Vol. 71, No. 4, May 1966, pp. 248-263.

[14] P. Kokotovic, H.K. Khalil, J. O'Reilly *Singular Perturbation Methods in Control*, SIAM, Philadelphia, 1999.

[15] A.E. Petropoulos, J.A. Sims, "A Review of Some Exact Solutions to the Planar Equations of Motion of Thrusting Spacecraft", *2nd International Symposium Low Thrust Trajectories*, Toulouse, France, 8th June 2002.

[16] I.M. Ross, Q. Gong, P. Sekhavat, "Low-Thrust, High Accuracy Trajectory Optimization", *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 4, July-August 2007, pp. 921-933.

[17] H.0. Ruppe, *Introduction to Astronautics* (2 volumes), Academic Press, New York, 1966.

[18] D.J. Scheeres, J.S. Hudson, "Reduction of Low-Thrust Continuous Controls for Trajectory Dynamics", *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 3, May-June 2009, pp. 780-787.

[19] D. Sharma, D.J. Scheeres, "Solar System Escape Trajectories Using Solar Sails", *Journal of Spacecraft and Rockets*, Vol.41, No. 4, 2004, pp. 684-687.

[20] M. Vasile, C. Colombo, G. Radice, "Semi-Analitical Solution for the Optimal Low-Thrust Deflection of Near-Earth Objects", *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 3, May-June 2009, pp. 796-809.

[21] M. Vasile, "Direct Transcription by FET for Optimal Space Trajectory Design", *Internal Report*, DIA-SR 99-02.

[22] M. Vasile, F. Bernelli-Zazzera, "Optimizing Low-Thrust and Gravity Assist Maneuvers to Design Interplanetary Trajectories", *Journal of the Astronautical Sciences*, Volume 51, n.1, January-March 2003.

[23] M. Vasile, F. Bernelli-Zazzera, "Targeting a Heliocentric Orbit Combining Low-Thrust Propulsion and Gravity Assist Manoeuvres", *16th International Symposium on Space Flight Dynamics*, Pasadena, 3-7 December 2001.

[24] M. Vasile, S. Campagnola, "Design of Low-Thrust Multi-Gravity Assist Trajectories to Europa", *Journal of the British Interplanetary Society*, Vol. 62, No. 1, January 2009, pp. 15-31.

[25] B.J. Wall, B.A. Conway, "Shape-Based Approach to Low-Thrust Rendezvous Trajectory Design", *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 1, January-February 2009, pp. 95-101.

[26] J.R. Wertz, *Spacecraft Attitude Determination and Control*, Kluwert Academic Publishers, Dordecht, 1978.