2 Gauge Freedom in Astrodynamics

MICHAEL EFROIMSKY
US Naval Observatory

Contents

2.1 Introduction

2.1.1 What this chapter is about

Both orbital and attitude dynamics employ the method of variation of parameters. In a non-perturbed setting, the coordinates (or the Euler angles) get expressed as functions of the time and six adjustable constants called elements. Under disturbance, each such expression becomes ansatz, the "constants" being endowed with time dependence. The perturbed velocity (linear or angular) consists of a partial time derivative and a convective term containing time derivatives of the "constants." It can be shown that this construction leaves one with a freedom to impose three arbitrary conditions upon the "constants" and/or their derivatives. Out of convenience, the Lagrange constraint is often imposed. It nullifies the convective term and thereby guarantees that under perturbation the functional dependence of the velocity upon the time and "constants" stays the same as in the undisturbed case. "Constants" obeying this condition are called osculating elements.

The "constants" chosen to be canonical are called Delaunay elements, in the orbital case, or Andoyer elements, in the spin case. (As some of the Andoyer elements are time-dependent even in the free-spin case, the role of "constants" is played by these elements' initial values.) The Andoyer and Delaunay sets of elements share a feature not readily apparent: in certain cases the standard equations render these elements non-osculating.

In orbital mechanics, elements calculated via the standard planetary equations come out non-osculating when perturbations depend on velocities. To keep elements osculating under such perturbations, the equations must be amended with extra terms that are not parts of the disturbing function [1, 2]. For the Kepler elements, this merely complicates the equations. In the case of Delaunay parameterisation, these extra terms not only complicate the equations, but also destroy their canonicity. So under velocity-dependent disturbances, osculation and canonicity are incompatible.

Similarly, in spin dynamics the Andoyer elements come out non-osculating under angular-velocity-dependent perturbation (a switch to a non-inertial frame being one such case). Amendment of the dynamical equations only with extra terms in the Hamiltonian makes the equations render non-osculating Andoyer elements. To make them osculating, more terms must enter the equations (and the equations will no longer be canonical).

It is often convenient to deliberately deviate from osculation by substituting the Lagrange constraint with an arbitrary condition that gives birth to a family of non-osculating elements. The freedom in choosing this condition is analogous to the gauge freedom. Calculations in non-osculating variables are mathematically valid and sometimes highly advantageous, but their physical interpretation is non-trivial. For example, non-osculating orbital elements parameterise instantaneous conics not tangent to the orbit, so the non-osculating inclination will be different from the real inclination of the physical orbit.

We present examples of situations in which ignoring of the gauge freedom (and of the unwanted loss of osculation) leads to oversights.

2.1.2 *Historical prelude*

The orbital dynamics is based on the variation-of-parameters method, invention whereof is attributed to Euler [3, 4] and Lagrange [5–9]. Though both greatly contributed to this approach, its initial sketch was offered circa 1687 by Newton in his unpublished *Portsmouth Papers*. Very succinctly, Newton brought up this issue also in Cor. 3 and 4 of Prop. 17 in the first book of his *Principia*.

Geometrically, the part and parcel of this method is representation of an orbit as a set of points, each of which is contributed by a member of some chosen family of curves $C(\kappa)$, where κ stands for a set of constants that number a particular curve within the family. (For example, a set of three constants $\kappa = \{a, b, c\}$ defines one particular hyperbola $y = ax^2 + bx + c$ out of many). This situation is depicted on Fig. 2.1. Point A of the orbit coincides with some point λ_1 on a curve $C(\kappa_1)$. Point B of the orbit coincides with point λ_2 on some other curve $C(\kappa_2)$ of the same family, etc. This way, orbital motion from A to B becomes a superposition of motion along C_κ from λ_1 to λ_2 and a gradual distortion of the curve C_κ from the shape $C(\kappa_1)$ to the shape $C(\kappa_2)$. In a loose language, the motion along the orbit consists of steps along an instantaneous curve $C(\kappa)$ which itself is evolving while those steps are being made.

Normally, the family of curves C_κ is chosen to be that of ellipses or that of hyperbolae, κ being six orbital elements, and λ being the time. However, if we disembody this idea of its customary implementation, we shall see that it is of a far more general nature and contains three aspects:

1. A trajectory may be assembled of points contributed by a family of curves of an essentially arbitrary type, not just conics.
2. It is not necessary to choose the family of curves tangent to the orbit. As we shall see below, it is often beneficial to choose those non-tangent. We shall also see examples

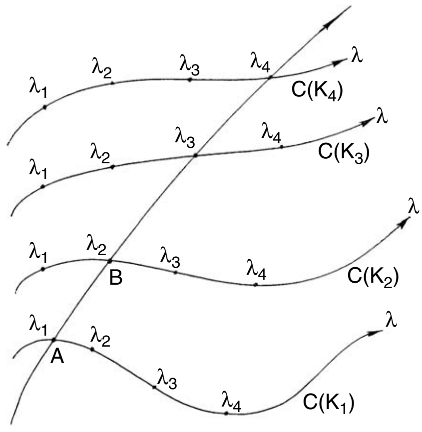

Fig. 2.1. Each point of the orbit is contributed by a member of some family of curves $C(\kappa)$ of a certain type, κ standing for a set of constants that number a particular curve within the family. Motion from A to B is, first, due to the motion along the curve $C(\kappa)$ from λ_1 to λ_2 and, second, due to the fact that during this motion the curve itself was evolving from $C(\kappa_1)$ to $C(\kappa_2)$.

when in orbital calculations this loss of tangentiality (loss of osculation) takes place and goes unnoticed.

3. The approach is general and can be applied, for example, to Euler's angles. A disturbed rotation can be thought of as a series of steps (small turns) along different Eulerian cones. An Eulerian cone is an orbit (on the Euler angles' manifold) corresponding to an unperturbed spin state. Just as a transition from one instantaneous Keplerian conic to another is caused by disturbing forces, so a transition from one instantaneous Eulerian cone to another is dictated by external torques or other perturbations. Thus, in the attitude mechanics, the Eulerian cones play the same role as the Keplerian conics do in the orbital dynamics. Most importantly, a perturbed rotation may be "assembled" of the Eulerian cones in an osculating or in a non-osculating manner. An unwanted loss of osculation in attitude mechanics happens in the same way as in the theory of orbits, but is much harder to notice. On the other hand, a deliberate choice of non-osculating rotational elements in attitude mechanics may sometimes be beneficial.

From the viewpoint of calculus, the concept of variation of parameters looks as follows. We have a system of differential equations to solve ("system in question") and a system of differential equations ("fiducial system") whose solution is known and contains arbitrary constants. We then use the known solution to the fiducial system as an ansatz for solving the system in question. The constants entering this ansatz are endowed with time dependence of their own, and the subsequent substitution of this known solution into the system in question yields equations for the "constants." The number of "constants" often exceeds that of equations in the system to solve. In this case we impose, by hand, arbitrary constraints upon the "constants." For example, in the case of a reduced N-body problem, we begin with $3(N-1)$ unconstrained second-order equations for $3(N-1)$ Cartesian coordinates. After a change of variables from the Cartesian coordinates to the orbital parameters, we end up with $3(N-1)$ differential equations for the $6(N-1)$ orbital variables. Evidently, $3(N-1)$ constraints are necessary.[1] To this end, the so-called Lagrange constraint (the condition of the instantaneous conics being tangent to the physical orbit) is introduced almost by default, because it is regarded natural. Two things should be mentioned in this regard:

First, what seems natural is not always optimal. The freedom of choice of the supplementary condition (the gauge freedom) gives birth to an internal symmetry (the gauge symmetry) of the problem. Most importantly, it can be exploited for simplifying the equations of motion for the "constants." On this issue we shall dwell in the current paper.

Second, the entire scheme may, in principle, be reversed and used to solve systems of differential equations with constraints. Suppose we have $N+M$ variables $C_j(t)$ obeying a system of N differential equations of the second order and M constraints expressed with first-order differential equations or with algebraic expressions. One possible approach to solving this system will be to assume that the variables C_j come about as constants emerging in a solution to some fiducial system of differential equations. Then our N second-order differential equations for $C_j(t)$ will be interpreted as a result of substitution of such an ansatz into the fiducial system with some perturbation, while our M constraints will be interpreted as weeding out of the redundant degrees of freedom. This subject is out of the scope of our paper and will not be developed here.

2.1.3 The simplest example of gauge freedom

Variation of constants first emerged in the non-linear context of celestial mechanics and later became a universal tool. We begin with a simple example offered in Newman and Efroimsky [10].

A harmonic oscillator disturbed by a force $\Delta F(t)$ gives birth to the initial-condition problem

$$\ddot{x}+x=\Delta F(t), \quad \text{with } x(0) \text{ and } \dot{x}(0) \text{ known}, \tag{2.1}$$

[1] In a fixed Cartesian frame, any solution to the unperturbed reduced 2-body problem can be written as

$$x_j = f_j(t, C_1, \ldots, C_6), \quad j = 1, 2, 3,$$

$$\dot{x}_j = g_j(t, C_1, \ldots, C_6), \quad g_j \equiv \left(\frac{\partial f_j}{\partial t}\right)_C$$

the adjustable constants C standing for orbital elements. Under disturbance, the solution is sought as

$$x_j = f_j(t, C_1(t), \ldots, C_6(t)), \quad j = 1, 2, 3,$$

$$\dot{x}_j = g_j(t, C_1(t), \ldots, C_6(t)) + \Phi_j(t, C_1(t), \ldots, C_6(t)), \quad g_j \equiv \left(\frac{\partial f_j}{\partial t}\right)_C, \quad \Phi_j \equiv \sum_r \frac{\partial f_j}{\partial C_r}\dot{C}_r.$$

Insertion of $x_j = f_j(t, C)$ into the perturbed gravity law yields three scalar equations for six functions $C_r(t)$. This necessitates imposition of three conditions upon C_r and $\dot{C}_r$. Under the simplest choice $\Phi_j = 0$, $j = 1, 2, 3$, the perturbed physical velocity $\dot{x}_j(t, C)$ has the same functional form as the unperturbed $g_j(t, C)$. Therefore, the instantaneous conics become tangent to the orbit (and the orbital elements $C_r(t)$ are called osculating).

whose solution may be sought using ansatz

$$x = C_1(t)\sin t + C_2(t)\cos t. \tag{2.2}$$

This will lead us to

$$\dot{x} = \left[\dot{C}_1(t)\sin t + \dot{C}_2(t)\cos t\right] + C_1(t)\cos t - C_2(t)\sin t. \tag{2.3}$$

It is common, at this point, to put the sum $\left[\dot{C}_1(t)\sin t + \dot{C}_2(t)\cos t\right]$ equal to zero, in order to remove the ambiguity stemming from the fact that we have only one equation for two variables. Imposition of this constraint is convenient but not obligatory. A more general way of fixing the ambiguity may be expressed as

$$\dot{C}_1(t)\sin t + \dot{C}_2(t)\cos t = \phi(t), \tag{2.4}$$

$\phi(t)$ being an arbitrary function of time. This entails:

$$\ddot{x} = \dot{\phi} + \dot{C}_1(t)\cos t - \dot{C}_2(t)\sin t - C_1(t)\sin t - C_2(t)\cos t, \tag{2.5}$$

summation whereof with Eq. (2.2) gives:

$$\ddot{x} + x = \dot{\phi} + \dot{C}_1(t)\cos t - \dot{C}_2(t)\sin t. \tag{2.6}$$

Substitution thereof into Eq. (2.1) yields the dynamical equation re-written in terms of the "constants" C_1, C_2. This equation, together with identity (2.4), will constitute the following system:

$$\begin{aligned} \dot{\phi} + \dot{C}_1(t)\cos t - \dot{C}_2(t)\sin t &= \Delta F(t), \\ \dot{C}_1(t)\sin t + \dot{C}_2(t)\cos t &= \phi(t), \end{aligned} \tag{2.7}$$

This leads to

$$\begin{aligned} \dot{C}_1 &= \Delta F\cos t - \frac{\mathrm{d}}{\mathrm{d}t}\left(\phi\cos t\right) \\ \dot{C}_2 &= -\Delta F\sin t + \frac{\mathrm{d}}{\mathrm{d}t}\left(\phi\sin t\right), \end{aligned} \tag{2.8}$$

the function $\phi(t)$ still remaining arbitrary.[2] Integration of Eq. (2.8) entails:

$$\begin{aligned} C_1 &= \int^t \Delta F\cos t'\mathrm{d}t' - \phi\cos t + a_1 \\ C_2 &= -\int^t \Delta F\sin t'\mathrm{d}t' + \phi\sin t + a_2. \end{aligned} \tag{2.9}$$

[2] Function $\phi(t)$ can afford being arbitrary, no matter what the initial conditions are to be. Indeed, for fixed $x(0)$ and $\dot{x}(0)$, the system $C_2(0) = x(0)$, $\phi(0) + C_1(0) = \dot{x}(0)$ solves for $C_1(0)$ and $C_2(0)$ for an arbitrary choice of $\phi(0)$.

Substitution of Eq. (2.9) into Eq. (2.2) leads to complete cancellation of the ϕ terms:

$$x = C_1 \sin t + C_2 \cos t = -\cos t \int^t \Delta F \sin t' \mathrm{d}t' + \sin t \int^t \Delta F \cos t' \mathrm{d}t' \\ + a_1 \sin t + a_2 \cos t \tag{2.10}$$

Naturally, the physical trajectory $x(t)$ remains invariant under the choice of gauge function $\phi(t)$, even though the mathematical description (2.9) of this motion in terms of the parameters C is gauge dependent. It is, however, crucial that a numerical solution of the system (2.8) will come out ϕ-dependent, because the numerical error will be sensitive to the choice of $\phi(t)$. This issue is now being studied by P. Gurfil and I. Klein [11], and the results are to be published soon.

It remains to notice that (2.8) is a simple analogue to the Lagrange-type system of planetary equations, system that, too, admits gauge freedom. (See subsection 2.2.2 below.)

2.1.4 Gauge freedom under a variation of the Lagrangian

The above example permits an evident extension [12, 13]. Suppose some mechanical system obeys the equation

$$\ddot{\boldsymbol{r}} = F(t, \boldsymbol{r}, \dot{\boldsymbol{r}}), \tag{2.11}$$

whose solution is known and has a functional form

$$\boldsymbol{r} = \boldsymbol{f}(t, C_1, \ldots, C_6), \tag{2.12}$$

C_j being adjustable constants to vary only under disturbance.

When a perturbation $\Delta \boldsymbol{F}$ gets switched on, the system becomes:

$$\ddot{\boldsymbol{r}} = \boldsymbol{F}(t, \boldsymbol{r}, \dot{\boldsymbol{r}}) + \Delta \boldsymbol{F}(t, \boldsymbol{r}, \dot{\boldsymbol{r}}), \tag{2.13}$$

and its solution will be sought in the form of

$$\boldsymbol{r} = \boldsymbol{f}(t, C_1(t), \ldots, C_6(t)). \tag{2.14}$$

Evidently,

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{f}}{\partial t} + \boldsymbol{\Phi}, \quad \boldsymbol{\Phi} \equiv \sum_{j=1}^{6} \frac{\partial \boldsymbol{f}}{\partial C_j} \dot{C}_j. \tag{2.15}$$

In defiance of what the textbooks advise, we do *not* put $\boldsymbol{\Phi}$ nil. Instead, we proceed further to

$$\ddot{\boldsymbol{r}} = \frac{\partial^2 \boldsymbol{f}}{\partial t^2} + \sum_{j=1}^{6} \frac{\partial^2 \boldsymbol{f}}{\partial t \, \partial C_j} \dot{C}_j + \dot{\boldsymbol{\Phi}}, \tag{2.16}$$

dot standing for a *full* time derivative. If we now insert the latter into the perturbed equation of motion (2.13) and if we recall that, according to (2.11),[3] $\partial^2 \boldsymbol{f}/\partial t^2 = \boldsymbol{F}$, then we shall obtain the equation of motion for the new variables $C_j(t)$:

$$\sum_{j=1}^{6} \frac{\partial^2 \boldsymbol{f}}{\partial t\, \partial C_j} \dot{C}_j + \dot{\boldsymbol{\Phi}} = \Delta \boldsymbol{F} \tag{2.17}$$

where

$$\boldsymbol{\Phi} \equiv \sum_{j=1}^{6} \frac{\partial \boldsymbol{f}}{\partial C_j} \dot{C}_j \tag{2.18}$$

so far is merely an identity. It will become a constraint after we choose a particular functional form $\boldsymbol{\Phi}\,(t;\, C_1, \ldots, C_6)$ for the gauge function $\boldsymbol{\Phi}$, i.e., if we choose that the sum $\sum \frac{\partial \boldsymbol{f}}{\partial C_j} \dot{C}_j$ be equal to some arbitrarily fixed function $\boldsymbol{\Phi}(t;\, C_1, \ldots, C_6)$ of the time and of the variable "constants." This arbitrariness exactly parallels the gauge invariance in electrodynamics: on the one hand, the choice of the functional form of $\boldsymbol{\Phi}(t;\, C_1, \ldots, C_6)$ will never[4] influence the eventual solution for the physical variable $\boldsymbol{r}$; on the other hand, though, a qualified choice may considerably simplify the process of finding the solution. To illustrate this, let us denote by $\mathbf{g}(t, C_1, \ldots, C_6)$ the functional dependence of the unperturbed velocity on the time and adjustable constants:

$$\mathbf{g}(t, C_1, \ldots, C_6) \equiv \frac{\partial}{\partial t} \boldsymbol{f}(t, C_1, \ldots, C_6), \tag{2.19}$$

and rewrite the above system as

$$\sum_j \frac{\partial \mathbf{g}}{\partial C_j} \dot{C}_j = -\dot{\boldsymbol{\Phi}} + \Delta \boldsymbol{F} \tag{2.20}$$

$$\sum_j \frac{\partial \boldsymbol{f}}{\partial C_j} \dot{C}_j = \boldsymbol{\Phi}. \tag{2.21}$$

If we now dot-multiply the first equation with $\partial \boldsymbol{f}/\partial C_i$ and the second one with $\partial \mathbf{g}/\partial C_i$, and then take the difference of the outcomes, we shall arrive at

$$\sum_j \left[C_n C_j \right] \dot{C}_j = (\Delta \boldsymbol{F} - \dot{\boldsymbol{\Phi}}) \cdot \frac{\partial \boldsymbol{f}}{\partial C_n} - \boldsymbol{\Phi} \cdot \frac{\partial \mathbf{g}}{\partial C_n}, \tag{2.22}$$

[3] We remind that in Eq. (2.11) there was no difference between a partial and a full time derivative, because at that point the integration "constants" C_i were indeed constant. Later, they acquired time dependence, and therefore the full time derivative implied in Eqs. (2.15–2.16) became different from the partial one implied in Eq. (2.11).

[4] Our usage of words "arbitrary" and "never" should be limited to the situations where the chosen gauge (2.21) does not contradict the equations of motion (2.20). This restriction, too, parallels a similar one present in field theories. Below we shall encounter a situation where this restriction becomes crucial.

the Lagrange brackets being defined in a gauge-invariant (i.e., $\mathbf{\Phi}$-independent) fashion.[5] If we agree that $\mathbf{\Phi}$ is a function of both the time and the parameters C_n, but not of their derivatives,[6] then the right-hand side of Eq. (2.22) will implicitly contain the first time derivatives of C_n. It will then be reasonable to move these to the left-hand side. Hence, Eq. (2.22) will be reshaped into

$$\sum_j \left([C_n C_j] + \frac{\partial \boldsymbol{f}}{\partial C_n} \cdot \frac{\partial \mathbf{\Phi}}{\partial C_j} \right) \frac{\mathrm{d}C_j}{\mathrm{d}t} = \frac{\partial \mathbf{f}}{\partial C_n} \cdot \Delta \boldsymbol{F} - \frac{\partial \mathbf{f}}{\partial C_n} \cdot \frac{\partial \mathbf{\Phi}}{\partial t} - \frac{\partial \mathbf{g}}{\partial C_n} \cdot \mathbf{\Phi}. \tag{2.23}$$

This is the general form of the gauge-invariant perturbation equations, that follows from the variation-of-parameters method applied to problem (2.13), for an arbitrary perturbation $\boldsymbol{F}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t)$ and under the simplifying assumption that the arbitrary gauge function $\mathbf{\Phi}$ is chosen to depend on the time and the parameters C_n, but not on their derivatives.[7] Assume that our problem (2.13) is not simply mathematical but is an equation of motion for some physical setting, so that $\boldsymbol{F}$ is a physical force corresponding to some undisturbed Lagrangian $\mathcal{L}_o$, and $\Delta \boldsymbol{F}$ is a force perturbation generated by a Lagrangian variation $\Delta\mathcal{L}$. If, for example, we begin with $\mathcal{L}_o(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \dot{\boldsymbol{r}}^2/2 - U(\boldsymbol{r}, t)$, momentum $\mathbf{p} = \dot{\boldsymbol{r}}$, and Hamiltonian $\mathcal{H}_o(\boldsymbol{r}, \mathbf{p}, t) = \mathbf{p}^2/2 + U(\boldsymbol{r}, t)$, then their disturbed counterparts will read:

$$\mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \frac{\dot{\boldsymbol{r}}^2}{2} - U(\boldsymbol{r}) + \Delta\mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t), \tag{2.24}$$

$$\boldsymbol{p} = \dot{\boldsymbol{r}} + \frac{\partial \Delta\mathcal{L}}{\partial \dot{\boldsymbol{r}}}, \tag{2.25}$$

$$\mathcal{H} = \mathbf{p}\dot{\boldsymbol{r}} - \mathcal{L} = \frac{\mathbf{p}^2}{2} + U + \Delta\mathcal{H}, \tag{2.26}$$

$$\Delta\mathcal{H} \equiv -\Delta\mathcal{L} - \frac{1}{2}\left(\frac{\partial \Delta\mathcal{L}}{\partial \dot{\boldsymbol{r}}}\right)^2. \tag{2.27}$$

---

[5] The Lagrange-bracket matrix is defined in a gauge-invariant way:

$$\sum_j [C_n C_j] \equiv \frac{\partial \boldsymbol{f}}{\partial C_n} \cdot \frac{\partial \mathbf{g}}{\partial C_j} - \frac{\partial \vec{\mathbf{g}}}{\partial C_n} \cdot \frac{\partial \boldsymbol{f}}{\partial C_j}.$$

and so is its inverse, the matrix composed of the Poisson brackets

$$\{C_n C_j\} \equiv \frac{\partial C_n}{\partial \boldsymbol{f}} \cdot \frac{\partial C_j}{\partial \mathbf{g}} - \frac{\partial C_n}{\partial \mathbf{g}} \cdot \frac{\partial C_j}{\partial \boldsymbol{f}}.$$

Evidently, Eq. (2.22) yields

$$\dot{C}_n = \sum_j \{C_n C_j\} \left[\frac{\partial \boldsymbol{f}}{\partial C_j} \cdot (\Delta \boldsymbol{F} - \dot{\mathbf{\Phi}}) - \mathbf{\Phi} \cdot \frac{\partial \mathbf{g}}{\partial C_j} \right].$$

[6] The necessity to fix a functional form of $\mathbf{\Phi}(t; C_1, \ldots, C_6)$, i.e., to impose three arbitrary conditions upon the "constants" C_j, evidently follows from the fact that, on the one hand, in the ansatz (2.14) we have six variables $C_n(t)$ and, on the other hand, the number of scalar equations of motion (i.e., Cartesian projections of the perturbed vector equation (2.13)) is only three. This necessity will become even more mathematically transparent after we cast the perturbed equation (2.13) into the normal form of Cauchy. (see Appendix1)

[7] We may also impart the gauge function with dependence upon the parameters' time derivatives of all orders. This will yield higher-than-first-order derivatives in Eq. (2.23). In order to close this system, one will then have to impose additional initial conditions, beyond those on $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$.

The Euler–Lagrange equation written for the perturbed Lagrangian (2.24) is:

$$\ddot{\mathbf{r}} = -\frac{\partial U}{\partial \boldsymbol{r}} + \Delta \boldsymbol{F}, \tag{2.28}$$

where the disturbing force is given by

$$\Delta \boldsymbol{F} \equiv \frac{\partial \Delta \mathcal{L}}{\partial \boldsymbol{r}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}\right). \tag{2.29}$$

Its substitution in Eq. (2.23) yields the generic form of the equations in terms of the Lagrangian disturbance [2]:

$$\begin{aligned}\sum_j \left([C_n C_j] + \frac{\partial \boldsymbol{f}}{\partial C_n} \cdot \frac{\partial}{\partial C_j}\left(\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}} + \boldsymbol{\Phi}\right)\right)\frac{\mathrm{d}C_j}{\mathrm{d}t} &= \frac{\partial}{\partial C_n}\left[\Delta \mathcal{L} + \frac{1}{2}\left(\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}\right)^2\right] \\ &\quad - \left(\frac{\partial \mathbf{g}}{\partial C_n} + \frac{\partial \boldsymbol{f}}{\partial C_n}\frac{\partial}{\partial t} + \frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}\frac{\partial}{\partial C_n}\right) \cdot \left(\boldsymbol{\Phi} + \frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}\right).\end{aligned} \tag{2.30}$$

This equation not only reveals the convenience of the special gauge

$$\boldsymbol{\Phi} = -\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}, \tag{2.31}$$

(which reduces to $\boldsymbol{\Phi} = 0$ in the case of velocity-independent perturbations), but also explicitly demonstrates how the Hamiltonian variation comes into play: it is easy to notice that, according to Eq. (2.27), the sum in square brackets on the right-hand side of Eq. (2.30) is equal to $-\Delta \mathcal{H}$, so the above equation takes the form $\sum_j [C_n\, C_j]\dot{C}_j = -\partial \Delta \mathcal{H}/\partial C_n$. All in all, it becomes clear that the trivial gauge, $\boldsymbol{\Phi} = 0$, leads to the maximal simplification of the variation-of-parameters equations expressed through the disturbing force: it follows from Eq. (2.22) that

$$\sum_j \left[C_n\, C_j\right]\dot{C}_j = \Delta \boldsymbol{F} \cdot \frac{\partial \boldsymbol{f}}{\partial C_n}, \text{ provided we have chosen } \boldsymbol{\Phi} = 0. \tag{2.32}$$

However, the choice of the special gauge (2.31) entails the maximal simplification of the variation-of-parameters equations when they are formulated via a variation of the Hamiltonian:

$$\sum_j [C_n\, C_j]\frac{\mathrm{d}C_j}{\mathrm{d}t} = -\frac{\partial \Delta \mathcal{H}}{\partial C_n}, \text{ provided we have chosen } \boldsymbol{\Phi} = -\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}. \tag{2.33}$$

It remains to spell out the already obvious fact that, in case the unperturbed force $\boldsymbol{F}$ is given by the Newton gravity law (i.e., when the undisturbed setting is the reduced two-body problem), then the variable "constants" C_n are merely the orbital elements parameterising a sequence of instantaneous conics out of which we "assemble" the perturbed trajectory through Eq. (2.14). When the conics' parameterisation is chosen to be via the Kepler or the Delaunay variables, then Eq. (2.30) yields the gauge-invariant version of the Lagrange-type or the Delaunay-type planetary equations, accordingly. Similarly, Eq. (2.22) implements the gauge-invariant generalisation of the planetary equations in the Euler–Gauss form.

From Eq. (2.22) we see that the Euler–Gauss-type planetary equations will always assume their simplest form (2.32) under the gauge choice $\mathbf{\Phi} = 0$. In astronomy this choice is called "the Lagrange constraint." It makes the orbital elements osculating, i.e., guarantees that the instantaneous conics, parameterised by these elements, are tangent to the perturbed orbit.

From Eq. (2.33) one can easily notice that the Lagrange- and Delaunay-type planetary equations simplify maximally under the condition (2.31). This condition coincides with the Lagrange constraint $\mathbf{\Phi} = 0$ when the perturbation depends only upon positions (not upon velocities or momenta). Otherwise, condition (2.31) deviates from that of Lagrange, and the orbital elements rendered by Eq. (2.33) are no longer osculating (so that the corresponding instantaneous conics are no longer tangent to the physical trajectory).

Of an even greater importance will be the following observation. If we have a velocity-dependent perturbing force, we can always find the appropriate Lagrangian variation and, therefrom, the corresponding variation of the Hamiltonian. If now we simply add the negative of this Hamiltonian variation to the disturbing function, then the resulting Eq. (2.33) will render not the osculating elements but orbital elements of a different type, ones satisfying the non-Lagrange constraint (2.31). Since the instantaneous conics, parameterised by such non-osculating elements, will not be tangent to the orbit, then physical interpretation of such elements may be non-trivial. Besides, they will return a velocity different from the physical one.[8] This pitfall is well-camouflaged and is easy to fall in.

These and other celestial-mechanics applications of the gauge freedom will be considered in detail in Section 2.2 below.

2.1.5 *Canonicity versus osculation*

One more relevant development will come from the theory of canonical perturbations. Suppose that in the absence of disturbances we start out with a system

$$\dot{q} = \frac{\partial \mathcal{H}^{(o)}}{\partial p}, \quad \dot{p} = -\frac{\partial \mathcal{H}^{(o)}}{\partial q}. \tag{2.34}$$

q and p being the Cartesian or polar coordinates and their conjugated momenta, in the orbital case, or the Euler angles and their momenta, in the rotation case. Then we switch, via a canonical transformation

$$q = f(Q, P, t), \quad p = \chi(Q, P, t) \tag{2.35}$$

to

$$\dot{Q} = \frac{\partial \mathcal{H}^{*}}{\partial P} = 0, \quad \dot{P} = -\frac{\partial \mathcal{H}^{*}}{\partial Q} = 0, \quad \mathcal{H}^{*} = 0, \tag{2.36}$$

[8] We mean that substitution of the values of these elements in $\mathbf{g}(t; C_1(t), \ldots, C_6(t))$ will not give the right velocity. The correct physical velocity will be given by $\dot{\boldsymbol{r}} = \mathbf{g} + \mathbf{\Phi}$.

where Q and P denote the set of Delaunay elements, in the orbital case, or the initial values of the Andoyer variables, in the case of rigid-body rotation.

This scheme relies on the fact that, for an unperturbed motion (i.e., for an unperturbed Keplerian conic, in an orbital case; or for an undisturbed Eulerian cone, in the spin case) a six-constant parameterisation may be chosen so that:

1. the parameters are constants and, at the same time, are canonical variables $\{Q, P\}$ with a zero Hamiltonian $\mathcal{H}^*(Q, P) = 0$;
2. for constant Q and P, the transformation equations (2.35) are mathematically equivalent to the dynamical equations (2.34).

Under perturbation, the "constants" Q, P begin to evolve so that, after their substitution into

$$q = f\,(Q(t), P(t), t)\,, \quad p = \chi(Q(t), P(t), t), \tag{2.37}$$

(f, χ being the same functions as in (2.35)), the resulting motion obeys the disturbed equations

$$\dot{q} = \frac{\partial(\mathcal{H}^{(o)} + \Delta\mathcal{H})}{\partial p}, \quad \dot{p} = -\frac{\partial(\mathcal{H}^{(o)} + \Delta\mathcal{H})}{\partial q}. \tag{2.38}$$

We also want our "constants" Q and P to remain canonical and to obey

$$\dot{Q} = \frac{\partial\,(\mathcal{H}^* + \Delta\mathcal{H}^*)}{\partial P}, \quad \dot{P} = -\frac{\partial\,(\mathcal{H}^* + \Delta\mathcal{H}^*)}{\partial Q} \tag{2.39}$$

where

$$\mathcal{H}^* = 0 \quad \text{and} \quad \Delta\mathcal{H}^*(Q, P, t) = \Delta\mathcal{H}\,(q(Q, P, t), p(Q, P, t), t)\,. \tag{2.40}$$

Above all, we wish the perturbed "constants" $C = Q, P$ (the Delaunay elements, in the orbital case; or the initial values of the Andoyer elements, in the spin case) to osculate. This means that we want the perturbed velocity to be expressed by the same function of $C_j(t)$ and t as the unperturbed velocity. Let us check when this is possible. The perturbed velocity is

$$\dot{q} = g + \Phi, \tag{2.41}$$

where

$$g(C(t), t) \equiv \frac{\partial q(C(t), t)}{\partial t} \tag{2.42}$$

is the functional expression for the unperturbed velocity, while

$$\Phi(C(t), t) \equiv \sum_{j=1}^{6} \frac{\partial q(C(t), t)}{\partial C_j}\,\dot{C}_j(t) \tag{2.43}$$

is the convective term. Since we chose the "constants" C_j to make canonical pairs (Q, P) obeying Eq. (2.39–2.40), then insertion of Eq. (2.39) into Eq. (2.43) will result in

$$\Phi = \sum_{n=1}^{3} \frac{\partial q}{\partial Q_n}\,\dot{Q}_n(t) + \sum_{n=1}^{3} \frac{\partial q}{\partial P_n}\,\dot{P}_n(t) = \frac{\partial \Delta\mathcal{H}(q, p)}{\partial p}. \tag{2.44}$$

So canonicity is incompatible with osculation when $\Delta\mathcal{H}$ depends on p. Our desire to keep the perturbed equations (2.39) canonical makes the orbital elements Q, P non-osculating in a particular manner prescribed by Eq. (2.44). This breaking of gauge invariance reveals that the canonical description is marked with "gauge stiffness" (term suggested by Peter Goldreich).

We see that, under a momentum-dependent perturbation, we still can use the ansatz (2.37) for calculation of the coordinates and momenta, but can no longer use $\dot{q} = \partial q/\partial t$ for calculating the velocities. Instead, we must use $\dot{q} = \partial q/\partial t + \partial\Delta\mathcal{H}/\partial p$, and the elements C_j will no longer be osculating. In the case of orbital motion (when C_j are the non-osculating Delaunay elements), this will mean that the instantaneous ellipses or hyperbolae parameterised by these elements will not be tangent to the orbit [1]. In the case of spin, the situation will be similar, except that, instead of instantaneous Keplerian conics, one will be dealing with instantaneous Eulerian cones—a set of trajectories on the Euler-angles manifold, each of which corresponds to some non-perturbed spin state [14].

The main conclusion to be derived from this example is the following: whenever we encounter a disturbance that depends not only upon positions but also upon velocities or momenta, implementation of the afore described canonical-perturbation method necessarily yields equations that render non-osculating canonical elements. It is possible to keep the elements osculating, but only at the cost of sacrificing canonicity. For example, under velocity-dependent orbital perturbations (like inertial forces, or atmospheric drag, or relativistic correction) the equations for osculating Delaunay elements will no longer be Hamiltonian [12, 13].

Above in this subsection we discussed the disturbed velocity $\dot{q}$. How about the disturbed momentum? For sufficiently simple unperturbed Hamiltonians, it can be written down very easily. For example, for $\mathcal{H} = \mathcal{H}_o + \Delta\mathcal{H} = p^2/2m + U(q) + \Delta\mathcal{H}$ we get:

$$p = \dot{q} + \frac{\partial\Delta\mathcal{L}}{\partial\dot{q}} = g + \Phi + \frac{\partial\Delta\mathcal{L}}{\partial\dot{q}} = g + \left(\Phi - \frac{\partial\Delta\mathcal{H}}{\partial\dot{q}}\right) = g. \tag{2.45}$$

In this case, the perturbed momentum p coincides with the unperturbed one, g. In application to the orbital motion, this means that contact elements (i.e., the non-osculating orbital elements obeying Eq. (2.31)), when substituted in $g(t; C_1, \ldots, C_6)$, furnish not the correct perturbed velocity but the correct perturbed momentum, i.e., they osculate the orbit *in phase space*. That such elements must exist was pointed out long ago by Goldreich [15] and Brumberg et al. [16], though these authors did not study their properties in detail.

2.2 Gauge freedom in the theory of orbits

2.2.1 Geometrical meaning of the arbitrary gauge function $\mathbf{\Phi}$

As explained above, the content of subsection 2.1.4 becomes merely a formulation of the Lagrange theory of orbits, provided $\boldsymbol{F}$ stands for the Newton gravity force, so that the undisturbed setting is the two-body problem. Then Eq. (2.22) expresses the gauge-invariant (i.e., taken with an arbitrary gauge $\mathbf{\Phi}(t; C_1, \ldots, C_6)$) planetary equations in the Euler–Gauss form. These equations render orbital elements that are, generally, not osculating. Equation (2.32) stands for the customary Euler–Gauss-type system for osculating (i.e., obeying $\mathbf{\Phi} = 0$) orbital elements.

Similarly, Eq. (2.30) stands for the gauge-invariant Lagrange-type or Delaunay-type (dependent upon whether C_i stand for the Kepler or Delaunay variables) equations. Such equations yield elements, which, generally, are not osculating. In those equations, one could fix the gauge by putting $\mathbf{\Phi} = 0$, thus making the resulting orbital elements osculating. However, this would be advantageous only in the case of velocity-independent $\Delta\mathcal{L}$. Otherwise, a maximal simplification is achieved through a deliberate refusal from osculation: by choosing the gauge as in Eq. (2.31) one ends up with simple equations (2.33). Thus, gauge (2.31) simplifies the planetary equations. (See Eqs. (2.46–2.57) below.) Besides, in the case when the Delaunay parameterisation is employed, this gauge makes the equations for the Delaunay variables canonical for reasons explained above in subsection 2.1.4.

The geometrical meaning of the convective term $\mathbf{\Phi}$ becomes evident if we recall that a perturbed orbit is assembled of points, each of which is donated by one representative of a sequence of conics, as on Fig. 2.2 and Fig. 2.3 where the "walk" over the instantaneous conics may be undertaken either in a non-osculating manner or in the osculating manner. The physical velocity $\dot{\boldsymbol{r}}$ is always tangent to the perturbed orbit, while the unperturbed Keplerian velocity $\mathbf{g} \equiv \partial \boldsymbol{f}/\partial t$ is tangent to the instantaneous conic. Their difference is

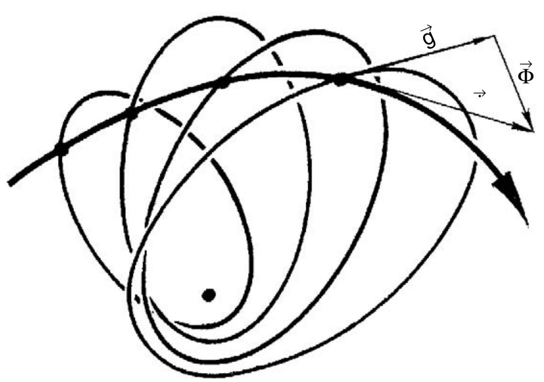

Fig. 2.2. The orbit is a set of points, each of which is donated by one of the confocal instantaneous ellipses that are not supposed to be tangent or even coplanar to the orbit. As a result, the physical velocity $\dot{\boldsymbol{r}}$ (tangent to the orbit) differs from the unperturbed Keplerian velocity $\mathbf{g}$ (tangent to the ellipse). To parameterise the depicted sequence of non-osculating ellipses, and to single it out of the other sequences, it is suitable to employ the difference between $\dot{\boldsymbol{r}}$ and $\mathbf{g}$, expressed as a function of time and six (non-osculating) orbital elements: $\mathbf{\Phi}(t, C_1, \ldots, C_6) = \dot{\boldsymbol{r}}(t, C_1, \ldots, C_6) - \mathbf{g}(t, C_1, \ldots, C_6)$. Evidently,

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{r}}{\partial t} + \sum_{j=1}^{6} \frac{\partial C_j}{\partial t} \dot{C}_j = \mathbf{g} + \mathbf{\Phi},$$

where the unperturbed Keplerian velocity is $\vec{\mathbf{g}} \equiv \partial \boldsymbol{r}/\partial t$. The convective term, which emerges under perturbation, is $\mathbf{\Phi} \equiv \sum (\partial \boldsymbol{r}/\partial C_j)\dot{C}_j$. When a particular functional dependence of $\mathbf{\Phi}$ on time and the elements is fixed, this function, $\mathbf{\Phi}(t, C_1, \ldots, C_6)$, is called gauge function or gauge velocity or, simply, gauge.

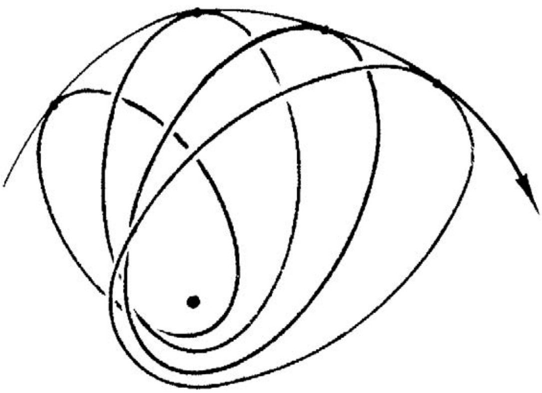

Fig. 2.3. The orbit is represented by a sequence of confocal instantaneous ellipses that are tangent to the orbit, i.e., osculating. Now, the physical velocity $\dot{\boldsymbol{r}}$ (tangent to the orbit) coincides with the unperturbed Keplerian velocity $\vec{\mathbf{g}}$ (tangent to the ellipse), so that their difference $\boldsymbol{\Phi}$ vanishes everywhere:

$$\boldsymbol{\Phi}(t, C_1, \ldots, C_6) \equiv \dot{\boldsymbol{r}}(t, C_1, \ldots, C_6) - \mathbf{g}(t, C_1, \ldots, C_6) = \sum_{j=1}^{6} \frac{\partial C_j}{\partial t} \dot{C}_j = 0.$$

This equality, called Lagrange constraint or Lagrange gauge, is the necessary and sufficient condition of osculation.

the convective term $\boldsymbol{\Phi}$. So if we use non-osculating orbital elements, then insertion of their values in $\boldsymbol{f}(t; C_1, \ldots, C_6)$ will yield a correct position of the body. However, their insertion in $\mathbf{g}(t; C_1, \ldots, C_6)$ will *not* give the right velocity. To get the correct velocity, one will have to add $\boldsymbol{\Phi}$. (See Appendix 1 for a more formal mathematical treatment in the normal form of Cauchy.)

When using non-osculating orbital elements, we must always be careful about their physical interpretation. On Fig. 2.2, the instantaneous conics are not supposed to be tangent to the orbit, nor are they supposed to be even coplanar thereto. (They may be even perpendicular to the orbit!—why not?) This means that, for example, the non-osculating element i may considerably differ from the real, physical inclination of the orbit.

We would add that the arbitrariness of choice of the function $\boldsymbol{\Phi}(t, C_1(t), \ldots, C_6(t))$ had been long known but never used in astronomy until a recent effort undertaken by several authors [1, 2, 10, 12, 13, 17, 18, 24] (Efroimsky 2005c). Substitution of the Lagrange constraint $\boldsymbol{\Phi} = 0$ with alternative choices does not influence the physical motion, but alters its mathematical description (i.e., renders different values of the orbital parameters $C_i(t)$). Such invariance of a physical picture under a change of parameterisation goes under the name of gauge freedom. It is a part and parcel of electrodynamics and other field theories. In mathematics, it is described in terms of fiber bundles. A clever choice of gauge often simplifies solution of the equations of motion. On the other hand, the gauge invariance may have implications upon numerical procedures. We mean the so-called "gauge drift," i.e., unwanted displacement in the gauge function $\boldsymbol{\Phi}$, caused by accumulation of numerical errors in the constants.

2.2.2 Gauge-invariant planetary equations of the Lagrange and Delaunay types

We present the gauge-invariant Lagrange-type equations, following Efroimsky and Goldreich [1]. These equations follow from (2.30) if we take into account the gauge-invariance (i.e., the $\mathbf{\Phi}$-independence) of the Lagrange-bracket matrix $[C_i\, C_j]$.

$$\begin{aligned}\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2}{na}\Bigg[&\frac{\partial(-\Delta\mathcal{H})}{\partial M_o} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial M_o}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial M_o}\\ &-\frac{\partial\boldsymbol{f}}{\partial M_o}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg],\end{aligned} \tag{2.46}$$

$$\begin{aligned}\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{1-e^2}{na^2e}\Bigg[&\frac{\partial\,(-\Delta\mathcal{H})}{\partial M_o} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial a}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial M_o}\\ &-\frac{\partial\boldsymbol{f}}{\partial M_o}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg] - \frac{(1-e^2)^{1/2}}{na^2e}\Bigg[\frac{\partial\,(-\Delta\mathcal{H})}{\partial\omega} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial\omega}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\\ &-\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial\omega} - \frac{\partial\boldsymbol{f}}{\partial\omega}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg],\end{aligned} \tag{2.47}$$

$$\begin{aligned}\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{-\cos i}{na^2(1-e^2)^{1/2}\sin i}\Bigg[&\frac{\partial\,(-\Delta\mathcal{H})}{\partial i} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial i}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial i}\\ &-\frac{\partial\boldsymbol{f}}{\partial i}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg] + \frac{(1-e^2)^{1/2}}{na^2e}\Bigg[\frac{\partial\,(-\Delta\mathcal{H})}{\partial e} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial e}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\\ &-\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial e} - \frac{\partial\boldsymbol{f}}{\partial e}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg],\end{aligned} \tag{2.48}$$

$$\begin{aligned}\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{\cos i}{na^2(1-e^2)^{1/2}\sin i}\Bigg[&\frac{\partial\,(-\Delta\mathcal{H})}{\partial\omega} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial\omega}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial\omega}\\ &-\frac{\partial\boldsymbol{f}}{\partial\omega}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg] - \frac{1}{na^2(1-e^2)^{1/2}\sin i}\Bigg[\frac{\partial\,(-\Delta\mathcal{H})}{\partial\Omega} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial\Omega}\left(\Phi+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\\ &-\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial\Omega} - \frac{\partial\boldsymbol{f}}{\partial\Omega}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg],\end{aligned} \tag{2.49}$$

$$\begin{aligned}\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \frac{1}{na^2(1-e^2)^{1/2}\sin i}\Bigg[&\frac{\partial\,(-\Delta\mathcal{H})}{\partial i} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial i}\left(\Phi+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial i}\\ &-\frac{\partial\boldsymbol{f}}{\partial i}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\Bigg],\end{aligned} \tag{2.50}$$

$$\frac{\mathrm{d}M_o}{\mathrm{d}t} = -\frac{1-e^2}{na^2e}\left[\frac{\partial(-\Delta\mathcal{H})}{\partial e} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial e}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial e}\right.$$
$$\left. -\frac{\partial\boldsymbol{f}}{\partial e}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\right] - \frac{2}{na}\left[\frac{\partial(-\Delta\mathcal{H})}{\partial a} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial a}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\right.$$
$$\left. -\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial a} - \frac{\partial\boldsymbol{f}}{\partial a}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\right]. \tag{2.51}$$

Similarly, the gauge-invariant Delaunay-type system can be written down as:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial(-\Delta\mathcal{H})}{\partial M_o} - \frac{\partial\Delta L}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial M_o}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial M_o}$$
$$-\frac{\partial\boldsymbol{r}}{\partial M_o}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right), \tag{2.52}$$

$$\frac{\mathrm{d}M_o}{\mathrm{d}t} = -\frac{\partial(-\Delta\mathcal{H})}{\partial L} + \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial L}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) + \left(\vec{\boldsymbol{\Phi}}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial L}$$
$$+\frac{\partial\boldsymbol{r}}{\partial L}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right), \tag{2.53}$$

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \frac{\partial(-\Delta\mathcal{H})}{\partial\omega} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial\omega}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial\omega}$$
$$-\frac{\partial\boldsymbol{r}}{\partial\omega}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right), \tag{2.54}$$

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = -\frac{\partial(-\Delta\mathcal{H})}{\partial G} + \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial G}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) + \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial G}$$
$$+\frac{\partial\boldsymbol{r}}{\partial G}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right), \tag{2.55}$$

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial(-\Delta\mathcal{H})}{\partial\Omega} - \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial\Omega}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) - \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial\Omega}$$
$$-\frac{\partial\boldsymbol{f}}{\partial\Omega}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right), \tag{2.56}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = -\frac{\partial(-\Delta\mathcal{H})}{\partial H} + \frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\frac{\partial}{\partial H}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right) + \left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)\frac{\partial\mathbf{g}}{\partial H}$$
$$+\frac{\partial\boldsymbol{r}}{\partial H}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{\Phi}+\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right). \tag{2.57}$$

where μ stands for the reduced mass, while

$$L \equiv \mu^{1/2} a^{1/2}, \quad G \equiv \mu^{1/2} a^{1/2} \left(1 - e^2\right)^{1/2}, \quad H \equiv \mu^{1/2} a^{1/2} \left(1 - e^2\right)^{1/2} \cos i. \tag{2.58}$$

The symbols $\mathbf{\Phi}, \boldsymbol{f}, \mathbf{g}$ now denote the functional dependencies of the gauge, position, and velocity upon the Delaunay, not Keplerian elements, and therefore these are functions different from $\mathbf{\Phi}, \boldsymbol{f}, \mathbf{g}$ used in Eqs. (2.46–2.51) where they stood for the dependencies upon the Kepler elements. (In Efroimsky [12, 13] the dependencies $\mathbf{\Phi}, \boldsymbol{f}, \mathbf{g}$ upon the Delaunay variables were equipped with tilde, to distinguish them from the dependencies upon the Kepler coordinates.)

To employ the gauge-invariant equations in analytical calculations is a delicate task: one should always keep in mind that, in case $\mathbf{\Phi}$ is chosen to depend not only upon time but also upon the "constants" (but not upon their derivatives), the right-hand sides of these equations will implicitly contain the first derivatives $\mathrm{d}C_i/\mathrm{d}t$, and one will have to move them to the left-hand sides (like in the transition from Eq. (2.22) to (2.23)). The choices $\mathbf{\Phi} = 0$ and $\mathbf{\Phi} = -\partial \Delta \mathcal{L}/\partial \dot{\boldsymbol{r}}$ are exceptions. (The most general exceptional gauge reads as $\mathbf{\Phi} = -\partial \Delta \mathcal{L}/\partial \dot{\boldsymbol{r}} + \eta(t)$, where $\eta(t)$ is an arbitrary function of time.)

As was expected from (2.30), both the Lagrange and Delaunay systems simplify in the gauge (2.31). Since for orbital motions we have $\partial \mathcal{H}/\partial \boldsymbol{p} = -\partial \Delta \mathcal{L}/\partial \dot{\boldsymbol{r}}$, then (2.31) coincides with Eq. (2.44). Hence, the Hamiltonian analysis (2.34–2.44) explains why it is exactly in the gauge (2.31) that the Delaunay system becomes symplectic. In physicists' parlance, the canonicity condition breaks the gauge symmetry by stiffly fixing the gauge (2.44), gauge that is equivalent, in the orbital case, to (2.31)—phenomenon called "gauge stiffness." The phenomenon may be looked upon also from a different angle. Above we emphasized that the gauge freedom implies essential arbitrariness in our choice of the functional form of $\mathbf{\Phi}(t; C_1, \ldots, C_6)$, provided the choice does not come into a contradiction with the equations of motion—an important clause that shows its relevance in the Delaunay-type Eqs. (2.52–2.57): we see that, for example, the Lagrange choice $\mathbf{\Phi} = 0$ (as well as any other choice different from Eq. (2.31)) is incompatible with the canonical structure of the equations of motion for the elements.

2.3 A practical example on gauges: a satellite orbiting a precessing oblate planet

Above we presented the Lagrange- and Delaunay-type planetary equations in the gauge-invariant form (i.e., for an arbitrary choice of the gauge function $\mathbf{\Phi}(t; C_1, \ldots, C_6)$) and for a generic perturbation $\Delta \mathcal{L}$ that may depend not only upon positions but also upon velocities and the time. We saw that the disturbing function is the negative Hamiltonian variation (which differs from the Lagrangian variation when the perturbation depends on velocities). Below, we shall also see that the functional dependence of $\Delta \mathcal{H}$ upon the orbital elements is gauge-dependent.

2.3.1 The gauge freedom and the freedom of frame choice

In the most compressed form, implementation of the variation-of-constants method in orbital mechanics looks like this. A generic solution to the two-body problem is expressed with

$$\boldsymbol{r} = \boldsymbol{f}\,(C, t)\,, \tag{2.59}$$

$$\left(\frac{\partial \boldsymbol{f}}{\partial t}\right)_C = \mathbf{g}\,(C, t)\,, \tag{2.60}$$

$$\left(\frac{\partial \mathbf{g}}{\partial t}\right)_C = -\frac{\mu}{f^2}\frac{\boldsymbol{f}}{f} \tag{2.61}$$

and is used as an ansatz to describe the perturbed motion:

$$\boldsymbol{r} = \boldsymbol{f}(C(t), t), \tag{2.62}$$

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{f}}{\partial t} + \frac{\partial \boldsymbol{f}}{\partial C_i}\frac{\mathrm{d}C_i}{\mathrm{d}t} = \mathbf{g} + \boldsymbol{\Phi}, \tag{2.63}$$

$$\ddot{\mathbf{r}} = \frac{\partial \mathbf{g}}{\partial t} + \frac{\partial \mathbf{g}}{\partial C_i}\frac{\mathrm{d}C_i}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = -\frac{\mu}{f^2}\frac{\boldsymbol{f}}{f} + \frac{\partial \mathbf{g}}{\partial C_i}\frac{\mathrm{d}C_i}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t}. \tag{2.64}$$

As can be seen from Eq. (2.63), our choice of a particular gauge is equivalent to a particular way of decomposition of the physical motion into a movement with velocity $\mathbf{g}$ along the instantaneous conic, and a movement caused by the conic's deformation at the rate $\boldsymbol{\Phi}$. Beside the fact that we decouple the physical velocity $\dot{\boldsymbol{r}}$ in a certain proportion between these two movements, $\mathbf{g}$ and $\boldsymbol{\Phi}$, it also matters *what* physical velocity (i.e., velocity relative to what frame) is decoupled in this proportion. Thus, the choice of gauge does not exhaust all freedom: one can still choose *in what frame* to write ansatz (2.62)—one can write it in inertial axes or in some accelerated or/and rotating ones. For example, in the case of a satellite orbiting a precessing oblate primary it is most *convenient* to write the ansatz in a frame co-precessing (but not corotating with the planet's equator.

The kinematic formulae (2.62–2.64) do not yet contain information about our choice of the reference system wherein to employ variation of constants. This information shows up only when (2.62) and (2.64) get inserted into the equation of motion $\ddot{\boldsymbol{r}} + (\mu \boldsymbol{r}/r^3) = \Delta \boldsymbol{F}$ to render

$$\frac{\partial \mathbf{g}}{\partial C_i}\frac{\mathrm{d}C_i}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = \Delta \boldsymbol{f} = \frac{\partial \Delta \mathcal{L}}{\partial \boldsymbol{r}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \Delta \mathcal{L}}{\partial \dot{\boldsymbol{r}}}\right). \tag{2.65}$$

Information about the reference frame, where we employ the method and define the elements C_i, is contained in the expression for the perturbing force $\Delta \boldsymbol{F}$. If the operation is carried out in an inertial system, $\Delta \boldsymbol{F}$ contains only physical forces. If we work in a frame moving with a linear acceleration $\vec{\mathbf{a}}$, then $\Delta \boldsymbol{F}$ also contains the inertial force $-\vec{\mathbf{a}}$. In case this coordinate frame also rotates relative to inertial ones at a rate $\boldsymbol{\mu}$, then $\Delta \boldsymbol{F}$ also includes the inertial contributions $-2\boldsymbol{\mu} \times \dot{\boldsymbol{r}} - \dot{\boldsymbol{\mu}} \times \boldsymbol{r} - \boldsymbol{\mu} \times (\boldsymbol{\mu} \times \dot{\boldsymbol{r}})$. When studying orbits about an oblate precessing planet, it is most convenient (though not obligatory) to apply the variation-of-parameters method in axes coprecessing with the planet's equator of date: it is in this coordinate system that one should write ansatz (2.62) and decompose $\dot{\boldsymbol{r}}$ into $\mathbf{g}$ and $\boldsymbol{\Phi}$. This convenient choice of coordinate system will still leave one with the freedom of gauge nomination: in the said coordinate system, one will still have to decide what function $\boldsymbol{\Phi}$ to insert in (2.63).

2.3.2 *The disturbing function in a frame co-precessing with the equator of date*

The equation of motion in the inertial frame is

$$\boldsymbol{r}'' = -\frac{\partial U}{\partial \boldsymbol{r}}, \tag{2.66}$$

where U is the total gravitational potential, and time derivatives in the inertial axes are denoted by primes. In a coordinate system precessing at angular rate $\boldsymbol{\mu}(t)$, Eq. (2.66) becomes:

$$\begin{aligned}\ddot{\mathbf{r}} &= -\frac{\partial U}{\partial \boldsymbol{r}} - 2\boldsymbol{\mu}\times\dot{\boldsymbol{r}} - \dot{\boldsymbol{\mu}}\times\boldsymbol{r} - \boldsymbol{\mu}\times(\boldsymbol{\mu}\times\boldsymbol{r})\\ &= -\frac{\partial U_o}{\partial \boldsymbol{r}} - \frac{\partial \Delta U}{\partial \boldsymbol{r}} - 2\boldsymbol{\mu}\times\dot{\boldsymbol{r}} - \dot{\boldsymbol{\mu}}\times\boldsymbol{r} - \boldsymbol{\mu}\times(\boldsymbol{\mu}\times\boldsymbol{r}),\end{aligned} \tag{2.67}$$

dots standing for time derivatives in the co-precessing frame, and $\boldsymbol{\mu}$ being the coordinate system's angular velocity relative to an inertial frame. Formula (2.125) in the Appendix gives the expression for $\boldsymbol{\mu}$ in terms of the longitude of the node and the inclination of the equator of date relative to that of epoch. The physical (i.e., not associated with inertial forces) potential $U(\boldsymbol{r})$ consists of the (reduced) two-body part $U_o(\boldsymbol{r}) \equiv -GM\boldsymbol{r}/r^3$ and a term $\Delta U(\boldsymbol{r})$ caused by the planet's oblateness (or, generally, by its triaxiality).

To implement variation of the orbital elements defined in this frame, we note that the disturbing force on the right-hand side of Eq. (2.67) is generated, according to Eq. (2.65), by

$$\Delta\mathcal{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = -\Delta U(\boldsymbol{r}) + \dot{\boldsymbol{r}}\cdot(\boldsymbol{\mu}\times\boldsymbol{r}) + \frac{1}{2}(\boldsymbol{\mu}\times\boldsymbol{r})\cdot(\boldsymbol{\mu}\times\boldsymbol{r}). \tag{2.68}$$

Since

$$\frac{\partial \Delta\mathcal{L}}{\partial \dot{\mathbf{r}}} = \boldsymbol{\mu}\times\boldsymbol{r}, \tag{2.69}$$

then

$$\boldsymbol{p} = \dot{\boldsymbol{r}} + \frac{\partial \Delta\mathcal{L}}{\partial \dot{\mathbf{r}}} = \dot{\boldsymbol{r}} + \boldsymbol{\mu}\times\boldsymbol{r} \tag{2.70}$$

and, therefore, the corresponding Hamiltonian perturbation reads:

$$\begin{aligned}\Delta\mathcal{H} &= -\left[\Delta\mathcal{L} + \frac{1}{2}\left(\frac{\partial \Delta\mathcal{L}}{\partial \dot{\mathbf{r}}}\right)^2\right] = -\left[-\Delta U + \boldsymbol{p}\cdot(\boldsymbol{\mu}\times\boldsymbol{r})\right]\\ &= -\left[-\Delta U + (\boldsymbol{r}\times\boldsymbol{p})\cdot\boldsymbol{\mu}\right] = \Delta U - \mathbf{J}\cdot\boldsymbol{\mu},\end{aligned} \tag{2.71}$$

with vector $\mathbf{J} \equiv \boldsymbol{r} \times \boldsymbol{p}$ being the satellite's orbital angular momentum in the inertial frame.

According to (2.63) and (2.70), the momentum can be written as

$$\boldsymbol{p} = \mathbf{g} + \boldsymbol{\Phi} + \boldsymbol{\mu} \times \boldsymbol{f}, \tag{2.72}$$

whence the Hamiltonian perturbation becomes

$$\Delta\mathcal{H} = -\left[\Delta\mathcal{L} + \frac{1}{2}\left(\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}}\right)^2\right] = -\left[-\Delta U + (\boldsymbol{f} \times \mathbf{g}) \cdot \boldsymbol{\mu} \right. \\ \left. + (\boldsymbol{\Phi} + \boldsymbol{\mu} \times \boldsymbol{f}) \cdot (\boldsymbol{\mu} \times \boldsymbol{f})\right]. \tag{2.73}$$

This is what one is supposed to plug in (2.30) or, the same, in (2.46–2.57).

2.3.3 *Planetary equations in a precessing frame, written in terms of contact elements*

In the subsection 2.3.2 we fixed our choice of the frame wherein to describe the orbit. By writing the Lagrangian and Hamiltonian variations as (2.68) and (2.73), we stated that our elements would be defined in the frame coprecessing with the equator. The frame being fixed, we are still left with the freedom of gauge choice. As evident from (2.33) or (2.46–2.57), the special gauge (2.31) ideally simplifies the planetary equations. Indeed, (2.31) and (2.69) together yield

$$\boldsymbol{\Phi} = -\frac{\partial\Delta\mathcal{L}}{\partial\dot{\boldsymbol{r}}} = -\boldsymbol{\mu} \times \boldsymbol{r} \equiv -\boldsymbol{\mu} \times \boldsymbol{f}, \tag{2.74}$$

wherefrom the Hamiltonian (2.73) becomes

$$\Delta\mathcal{H}^{(cont)} = -\left[-\Delta U(\boldsymbol{f}) + \boldsymbol{\mu} \cdot (\boldsymbol{f} \times \mathbf{g})\right], \tag{2.75}$$

while the planetary equations (2.30) get the shape

$$[C_r C_i] \frac{\mathrm{d}C_i}{\mathrm{d}t} = \frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial C_r}, \tag{2.76}$$

or, the same,

$$[C_r\, C_i] \frac{\mathrm{d}C_i}{\mathrm{d}t} = \frac{\partial}{\partial C_r}\left[-\Delta U(\boldsymbol{f}) + \boldsymbol{\mu} \cdot (\boldsymbol{f} \times \mathbf{g})\right], \tag{2.77}$$

where $\boldsymbol{f}$ and $\mathbf{g}$ stand for the undisturbed (two-body) functional expressions (2.59) and (2.60) of the position and velocity via the time and the chosen set of orbital elements. Planetary equations (2.76) were obtained with aid of (2.74), and therefore they render non-osculating orbital elements that are called contact elements. This is why we equipped the Hamiltonian (2.75) with superscript "(*cont*)." In distinction from the osculating elements, the contact ones osculate *in phase space*: (2.72) and (2.74) entail that $\boldsymbol{p} = \mathbf{g}$. As already mentioned in the end of section 2.1, existence of such elements was pointed out by Goldreich [15] and Brumberg et al. [16] long before the concept of gauge freedom was introduced in celestial mechanics. Brumberg et al. [16] simply *defined* these elements by the condition that their insertion in $\mathbf{g}(t; C_1, \ldots, C_6)$ returns not the perturbed velocity, but the perturbed momentum. Goldreich [15] defined these elements (without calling them "contact") differently. Having in mind inertial forces (2.67), he wrote down the corresponding Hamiltonian equation (2.71) and added its negative to the disturbing function of the standard planetary equations (without enriching the equations with any other terms). Then he noticed that those equations furnished non-osculating elements. Now we can easily see that both Goldreich's and Brumberg's definitions correspond to the gauge choice (2.31).

When one chooses the Keplerian parameterisation, then Eq. (2.77) becomes:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2}{na}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial M_o}, \tag{2.78}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{1-e^2}{na^2e}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial M_o} - \frac{(1-e^2)^{1/2}}{na^2e}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\omega}, \tag{2.79}$$

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{-\cos i}{na^2(1-e^2)^{1/2}\sin i}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial i} + \frac{(1-e^2)^{1/2}}{na^2e}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial e} \tag{2.80}$$

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{\cos i}{na^2(1-e^2)^{1/2}\sin i}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\omega} - \frac{1}{na^2(1-e^2)^{1/2}\sin i}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\Omega}, \tag{2.81}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \frac{1}{na^2(1-e^2)^{1/2}\sin i}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial i}, \tag{2.82}$$

$$\frac{\mathrm{d}M_o}{\mathrm{d}t} = -\frac{1-e^2}{na^2e}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial e} - \frac{2}{na}\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial a}. \tag{2.83}$$

The above equations implement an interesting pitfall. When describing orbital motion relative to a frame coprecessing with the equator of date, it is tempting to derive the Hamiltonian variation caused by the inertial forces, and to simply plug it, with a negative sign, into the disturbing function. This would entail equations (2.76–2.83) which, as demonstrated above, belong to the non-Lagrange gauge (2.31). The elements furnished by these equations are non-osculating, so that the conics parameterised by these elements are not tangent to the perturbed trajectory. For example, i gives the inclination of the instantaneous non-tangent conic, but differs from the real, physical physical (i.e., osculating), inclination of the orbit. This approach—when an inertial term is simply added to the disturbing function—was employed by Goldreich [15], Brumberg et al. [16], and Kinoshita [19], and many others. Goldreich and Brumberg noticed that this destroyed the osculation.

Goldreich [15] studied how the equinoctial precession of Mars influences the long-term evolution of Phobos' and Deimos' orbit inclinations. Goldreich assumed that the elements a and e stay constant; he also substituted the Hamiltonian variation (2.75) with its orbital average, which made his planetary equations render the secular parts of the elements. He assumed that the averaged physical term $\langle \Delta U \rangle$ is only due to the primary's oblateness:

$$\langle \Delta U \rangle = -\frac{n^2 J_2}{4} \rho^2 \frac{3\cos^2 i - 1}{(1-e^2)^{3/2}}, \tag{2.84}$$

ρ being the mean radius of the planet,[9] and n being the satellite's mean motion. To simplify the inertial term, Goldreich employed the well known formula

$$\boldsymbol{r} \times \mathbf{g} = \sqrt{Gma\,(1-e^2)}\,\boldsymbol{w}, \tag{2.85}$$

where

$$\boldsymbol{w} = \hat{\mathbf{x}}_1 \sin i \sin \Omega - \hat{\mathbf{x}}_2 \sin i \cos \Omega + \hat{\mathbf{x}}_3 \cos i \tag{2.86}$$

is a unit vector normal to the instantaneous ellipse, expressed through unit vectors $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{x}}_3$ associated with the co-precessing frame x_1, x_2, x_3 (axes x_1 and x_2 lying in the planet's equatorial plane of date, and x_1 pointing along the fiducial line wherefrom the longitude of the ascending node of the satellite orbit, Ω, is measured). This resulted in

$$\begin{aligned}\langle \Delta \mathcal{H}^{(cont)} \rangle = -\left[-\langle \Delta U \rangle + \langle \boldsymbol{\mu} \cdot (\boldsymbol{f} \times \mathbf{g}) \rangle\right] = -\frac{GmJ_2}{4} \frac{\rho_e^2}{a^3} \frac{3\cos^2 i - 1}{(1-e^2)^{3/2}} - \sqrt{Gma\,(1-e^2)} \\ (\mu_1 \sin i \sin \Omega - \mu_2 \sin i \cos \Omega + \mu_3 \cos i)\,,\end{aligned} \tag{2.87}$$

all letters now standing not for the appropriate variables but for their orbital averages. Substitution of this averaged Hamiltonian in (2.81–2.82) lead Goldreich, in assumption that both $|\dot{\boldsymbol{\mu}}| / \left(n^2 J_2 \sin i\right)$ and $|\boldsymbol{\mu}| / (n J_2 \sin i)$ are much less than unity, to the following system:

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} \approx -\frac{3}{2} n J_2 \left(\frac{\rho_e}{a}\right)^2 \frac{\cos i}{(1-e^2)^2}, \tag{2.88}$$

$$\frac{\mathrm{d}i}{\mathrm{d}t} \approx -\mu_1 \cos \Omega - \mu_2 \sin \Omega, \tag{2.89}$$

whose solution,

$$\begin{aligned} &i = -\frac{\mu_1}{\chi} \cos\left[-\chi\,(t-t_o) + \Omega_o\right] + \frac{\mu_2}{\chi} \sin\left[-\chi\,(t-t_o) + \Omega_o\right] + i_o, \\ &\Omega = -\chi\,(t-t_o) + \Omega_o \quad \text{where} \quad \chi \equiv \frac{3}{2} n J_2 \left(\frac{\rho_e}{a}\right)^2 \frac{\cos i}{(1-e^2)^2},\end{aligned} \tag{2.90}$$

[9] Goldreich used the non-sphericity parameter $J \equiv (3/2)\,(\rho_e/\rho)^2\,J_2$, where ρ_e is the mean *equatorial* radius.

tells us that in the course of equinoctial precession the satellite inclination oscillates about i_o.

Goldreich [15] noticed that his i and the other elements were not osculating, but he assumed that their secular parts would differ from those of the osculating ones only in the orders higher than $O(|\boldsymbol{\mu}|)$. Below we shall probe the applicability limits for this assumption. (See the end of subsection 2.3.5.)

2.3.4 *Planetary equations in a precessing frame, in terms of osculating elements*

When one introduces elements in the precessing frame and also demands that they osculate in this frame (i.e., obey the Lagrange constraint $\boldsymbol{\Phi} = 0$), the Hamiltonian variation reads:[10]

$$\Delta\mathcal{H}^{(osc)} = -\left[-\Delta U + \boldsymbol{\mu}\cdot(\boldsymbol{f}\times\mathbf{g}) + (\boldsymbol{\mu}\times\boldsymbol{f})\cdot(\boldsymbol{\mu}\times\boldsymbol{f})\right], \tag{2.91}$$

while Eq. (2.30) becomes:

$$\begin{aligned}[C_n C_i]\frac{\mathrm{d}C_i}{\mathrm{d}t} = &-\frac{\partial\Delta\mathcal{H}^{(osc)}}{\partial C_n} + \boldsymbol{\mu}\cdot\left(\frac{\partial\boldsymbol{f}}{\partial C_n}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial C_n}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial\boldsymbol{f}}{\partial C_n}\right)\\ &-(\boldsymbol{\mu}\times\boldsymbol{f})\frac{\partial}{\partial C_n}(\boldsymbol{\mu}\times\boldsymbol{f}).\end{aligned} \tag{2.92}$$

To ease the comparison of this equation with (2.77), it is convenient to split the expression (2.91) for $\Delta\mathcal{H}^{(osc)}$ into two parts:

$$\Delta\mathcal{H}^{(cont)} = -\left[R_{oblate}(\boldsymbol{f}, t) + \boldsymbol{\mu}\cdot(\boldsymbol{f}\times\mathbf{g})\right] \tag{2.93}$$

and

$$-(\boldsymbol{\mu}\times\boldsymbol{f})\cdot(\boldsymbol{\mu}\times\boldsymbol{f}), \tag{2.94}$$

and then to group the latter part with the last term on the right-hand side of (2.35):

$$\begin{aligned}[C_n C_i]\frac{\mathrm{d}C_i}{\mathrm{d}t} = &-\frac{\partial\Delta\mathcal{H}^{(cont)}}{\partial C_n} + \boldsymbol{\mu}\cdot\left(\frac{\partial\boldsymbol{f}}{\partial C_n}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial C_n}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial\boldsymbol{f}}{\partial C_n}\right)\\ &+(\boldsymbol{\mu}\times\boldsymbol{f})\frac{\partial}{\partial C_n}(\boldsymbol{\mu}\times\boldsymbol{f}).\end{aligned} \tag{2.95}$$

Comparison of this analytical theory with a straightforward numerical integration[11] has confirmed that the $O(|\boldsymbol{\mu}|^2)$ term in (2.95) may be neglected over time scales of, at

[10] Both $\Delta\mathcal{H}^{(cont)}$ and $\Delta\mathcal{H}^{(osc)}$ are equal to $-[-\Delta U(\boldsymbol{f}, t) + \boldsymbol{\mu}\cdot\mathbf{J}] = -[-\Delta U(\boldsymbol{f}, t) + \boldsymbol{\mu}\cdot(\boldsymbol{f}\times\boldsymbol{p})]$. However, the canonical momentum now is different from $\mathbf{g}$ and reads as: $\boldsymbol{p} = \mathbf{g} + (\boldsymbol{\mu}\times\boldsymbol{f})$. Hence, the functional forms of $\Delta\mathcal{H}^{(osc)}(\boldsymbol{f}, \boldsymbol{p})$ and $\Delta\mathcal{H}^{(can)}(\boldsymbol{f}, \boldsymbol{p})$ are different, though their values coincide.

[11] Credit for this comparison goes to Pini Gurfil and Valery Lainey.

least, hundreds of millions of years. In this approximation there is no difference between $\Delta\mathcal{H}^{(cont)}$ and $\Delta\mathcal{H}^{(osc)}$, so we shall write down the equations as:

$$[C_n\,C_i]\frac{\mathrm{d}C_i}{\mathrm{d}t} = -\frac{\partial\Delta\mathcal{H}^{(cont)}}{\partial C_n} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial C_n}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\vec{\mathbf{g}}}{\partial C_n}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial C_n}\right). \tag{2.96}$$

For C_i chosen as the Kepler elements, inversion of the Lagrange brackets in (2.90) will yield the following Lagrange-type system:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2}{na}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial M_o} - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial M_o}\right)\right], \tag{2.97}$$

$$\begin{aligned}\frac{\mathrm{d}e}{\mathrm{d}t} = {} & \frac{1-e^2}{na^2e}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial M_o} - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial M_o}\right)\right] - \frac{(1-e^2)^{1/2}}{na^2e}\\ & \times\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\omega} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial\omega}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial\omega}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial\omega}\right)\right],\end{aligned} \tag{2.98}$$

$$\begin{aligned}\frac{\mathrm{d}\omega}{\mathrm{d}t} = {} & \frac{-\cos i}{na^2(1-e^2)^{1/2}\sin i}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial i} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial i}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial i}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial i}\right)\right]\\ & + \frac{(1-e^2)^{1/2}}{na^2e}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial e} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial e}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial e}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial e}\right)\right],\end{aligned} \tag{2.99}$$

$$\begin{aligned}\frac{\mathrm{d}i}{\mathrm{d}t} = {} & \frac{\cos i}{na^2(1-e^2)^{1/2}\sin i}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\omega} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial\omega}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial\omega}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial\omega}\right)\right]\\ & - \frac{1}{na^2(1-e^2)^{1/2}\sin i}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial\Omega} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial\Omega}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial\Omega}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial\Omega}\right)\right],\end{aligned} \tag{2.100}$$

$$\begin{aligned}\frac{\mathrm{d}\Omega}{\mathrm{d}t} = {} & \frac{1}{na^2(1-e^2)^{1/2}\sin i}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial i} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial i}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\mathbf{g}}{\partial i}\right)\right.\\ & \left. - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial i}\right)\right],\end{aligned} \tag{2.101}$$

$$\frac{\mathrm{d}M_o}{\mathrm{d}t} = -\frac{1-e^2}{na^2e}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial e} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial e}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial \mathbf{g}}{\partial e}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial e}\right)\right]$$
$$-\frac{2}{na}\left[\frac{\partial\left(-\Delta\mathcal{H}^{(cont)}\right)}{\partial a} + \boldsymbol{\mu}\cdot\left(\frac{\partial \boldsymbol{f}}{\partial a}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial \mathbf{g}}{\partial a}\right) - \dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial \boldsymbol{f}}{\partial a}\right)\right], \tag{2.102}$$

terms $\boldsymbol{\mu}\cdot((\partial\boldsymbol{f}/\partial M_o)\times\mathbf{g} - (\partial\mathbf{g}/\partial M_o)\times\boldsymbol{f})$ being omitted in (2.97–2.98), because these terms vanish identically (see the Appendix to Efroimsky [14]).

2.3.5 Comparison of calculations performed in the two above gauges

Simply from looking at Eqs. (2.76–2.83) and (2.96–2.102) we notice that the difference in orbit descriptions performed in the two gauges emerges already in the first order of the precession rate $\boldsymbol{\mu}$ and in the first order of $\dot{\boldsymbol{\mu}}$.

Calculation of the $\boldsymbol{\mu}$- and $\dot{\boldsymbol{\mu}}$-dependent terms emerging in Eqs. (2.97–2.102) takes more than 20 pages of algebra. The resulting expressions are published in Efroimsky [20], their detailed derivation being available in web-archive preprint Efroimsky [20]. As an illustration, we present a couple of expressions:

$$\begin{aligned}
-\dot{\boldsymbol{\mu}}\cdot\left(\boldsymbol{f}\times\frac{\partial\boldsymbol{f}}{\partial i}\right) = a^2\frac{\left(1-e^2\right)^2}{(1+e\cos\nu)^2}\{&\dot{\mu}_1\left[-\cos\Omega\sin(\omega+\nu)\right.\\
&\left.-\sin\Omega\cos(\omega+\nu)\cos i\right]\sin(\omega+\nu)+\dot{\mu}_2\left[-\sin\Omega\sin(\omega+\nu)\right.\\
&\left.+\cos\Omega\cos(\omega+\nu)\cos i\right]\sin(\omega+\nu)\\
&+\dot{\mu}_3\sin(\omega+\nu)\cos(\omega+\nu)\sin i\},
\end{aligned} \tag{2.103}$$

$$\boldsymbol{\mu}\cdot\left(\frac{\partial\boldsymbol{f}}{\partial e}\times\mathbf{g} - \boldsymbol{f}\times\frac{\partial\vec{\mathbf{g}}}{\partial e}\right) = -\mu_\perp\frac{na^2\left(3e+2\cos\nu+e^2\cos\nu\right)}{(1+e\cos\nu)\sqrt{1-e^2}}, \tag{2.104}$$

ν denoting the true anomaly. The fact that almost none of these terms vanish reveals that Eqs. (2.76–2.83) and (2.96–2.102) may yield very different results, i.e., that the contact elements may differ from their osculating counterparts already in the first order of $\boldsymbol{\mu}$.

Luckily, in the practical situations we need not the elements *per se* but their secular parts. To calculate these, one can substitute both the Hamiltonian variation and the $\boldsymbol{\mu}$- and $\dot{\boldsymbol{\mu}}$-dependent terms with their orbital averages[12] calculated through

$$\langle\dots\rangle \equiv \frac{\left(1-e^2\right)^{3/2}}{2\pi}\int_0^{2\pi}\dots\frac{\mathrm{d}\nu}{(1+e\cos\nu)^2}. \tag{2.105}$$

[12] Mathematically, this procedure is, to say the least, not rigorous. In practical calculations it works well, at least over not too long time scales.

The situation might simplify very considerably if we could also assume that the precession rate $\boldsymbol{\mu}$ stays constant. Then in equations (2.97–2.102), we would assume $\boldsymbol{\mu} = constant$ and proceed with averaging the expressions $((\partial \boldsymbol{f}/\partial C_j) \times \mathbf{g} - \boldsymbol{f} \times (\partial \mathbf{g}/\partial C_j))$ only (while all the terms with $\dot{\boldsymbol{\mu}}$ will now vanish).

Averaging of the said terms is lengthy and is presented in the Appendix to Efroimsky [14]. All in all, we get, for constant $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} \cdot \left\langle \left(\frac{\partial \boldsymbol{f}}{\partial a} \times \vec{\mathbf{g}} - \boldsymbol{f} \times \frac{\partial \mathbf{g}}{\partial a} \right) \right\rangle = \boldsymbol{\mu} \cdot \left(\frac{\partial \boldsymbol{f}}{\partial a} \times \vec{\mathbf{g}} - \boldsymbol{f} \times \frac{\partial \mathbf{g}}{\partial a} \right) = \frac{3}{2} \mu_{\perp} \sqrt{\frac{Gm\,(1-e^2)}{a}}, \quad (2.106)$$

$$\boldsymbol{\mu} \cdot \left\langle \left(\frac{\partial \boldsymbol{f}}{\partial C_j} \times \mathbf{g} - \boldsymbol{f} \times \frac{\partial \mathbf{g}}{\partial C_j} \right) \right\rangle = 0, \qquad C_j = e, \Omega, \omega, i, M_o. \quad (2.107)$$

Since the orbital averages (2.107) vanish, then e will, along with a, stay constant for as long as our approximation remains valid. Besides, no trace of $\boldsymbol{\mu}$ will be left in the equations for Ω and i. This means that, in the assumed approximation and under the extra assumption of constant $\boldsymbol{\mu}$, the afore quoted analysis (2.84–2.90), offered by Goldreich [15], will remain valid at time scales which are not tool long.

In the realistic case of time-dependent precession, the averages of terms containing $\boldsymbol{\mu}$ and $\dot{\boldsymbol{\mu}}$ do not vanish (except for $\boldsymbol{\mu} \cdot ((\partial \boldsymbol{f}/\partial M_o) \times \mathbf{g} - \boldsymbol{f} \times (\partial \mathbf{g}/\partial M_o))$, which is identically nil). These terms show up in all equations (except in that for a) and influence the motion. Integration that includes these terms gives results very close to the Goldreich approximation (approximation (2.90) that neglects the said terms and approximates the secular parts of the non-osculating elements with those of their osculating counterparts). However, this agreement takes place only at time scales of order millions to dozens of millions of years. At larger time scales, differences begin to accumulate [21].

In real life, the equinoctial-precession rate of the planet, $\boldsymbol{\mu}$, is not constant. Since the equinoctial precession is caused by the solar torque acting on the oblate planet, this precession is regulated by the relative location and orientation of the Sun and the planetary equator. This is why $\boldsymbol{\mu}$ of a planet depends upon this planet's orbit precession caused by the pull from the other planets. This dependence is described by a simple model developed by Colombo [22].

2.4 Conclusions: how we benefit from the gauge freedom

In this chapter we gave a review of the gauge concept in orbital and attitude dynamics. Essentially, this is the freedom of choosing non-osculating orbital (or rotational) elements, i.e., the freedom of making them deviate from osculation in a known, prescribed, manner.

The advantage of elements introduced in a non-trivial gauge is that in certain situations the choice of such elements considerably simplifies the mathematical description of orbital and attitude problems. One example of such simplification is the Goldreich [15] approximation (2.90) for satellite orbiting a precessing oblate planet. Although performed in terms of non-osculating elements, Goldreich's calculation has the advantage of mathematical simplicity. Most importantly, later studies [20, 23] have confirmed that Goldreich's results, obtained for non-osculating elements, serves as a very good approximation for the osculating elements. To be more exact, the secular parts of these non-osculating elements coincide, in the first order over the precession-caused perturbation, with those of their osculating counterparts, the difference accumulating only at very long time scales—see the end of Section 2.3 above. A comprehensive investigation into this topic, with the relevant numerics, will be presented in Lainey et al. [21].

On the other hand, *neglect of the gauge freedom may sometimes produce camouflaged pitfalls caused by the fact that non-osculating elements lack evident physical meaning.* For example, the non-osculating "inclination" does not coincide with the real, physical inclination of the orbit. This happens because non-osculating elements parameterise instantaneous conics non-tangent to the orbit. Similar difficulties emerge in the theory of rigid-body rotation, when non-osculating Andoyer variables are employed.

Appendix 1. Mathematical formalities: Orbital dynamics in the normal form of Cauchy

Let us cast the perturbed equation

$$\ddot{\mathbf{r}} = \boldsymbol{F} + \Delta\boldsymbol{f} = -\frac{\mu}{r^2}\frac{\boldsymbol{r}}{r} + \Delta\boldsymbol{f} \tag{2.108}$$

into the normal form of Cauchy:

$$\dot{\boldsymbol{r}} = \mathbf{v}, \tag{2.109}$$

$$\dot{\mathbf{v}} = -\frac{\mu}{r^2}\frac{\boldsymbol{r}}{r} + \Delta\boldsymbol{f}\,(\boldsymbol{r}(t, C_1, \ldots, C_6), \mathbf{v}(t, C_1, \ldots, C_6), t)\,. \tag{2.110}$$

Insertion of our ansatz

$$\boldsymbol{r} = \boldsymbol{f}\,(t, C_1(t), \ldots, C_6(t))\,, \tag{2.111}$$

will make (2.109) equivalent to

$$\mathbf{v} = \frac{\partial \boldsymbol{f}}{\partial t} + \sum_i \frac{\partial \boldsymbol{f}}{\partial C_i}\dot{C}_i. \tag{2.112}$$

The function $\boldsymbol{f}$ is, by definition, the generic solution to the unperturbed equation

$$\ddot{\mathbf{r}} = \boldsymbol{F} = -\frac{\mu}{r^2}\frac{\boldsymbol{r}}{r}\cdot \tag{2.113}$$

This circumstance, along with (2.112), will transform (2.109) into

$$\sum_i \frac{\partial \mathbf{g}}{\partial C_i}\dot{C}_i + \dot{\mathbf{\Phi}} = \Delta\mathbf{F}\,(\boldsymbol{f}(t, C_1, \ldots, C_6),\ \mathbf{g}(t, C_1, \ldots, C_6) + \mathbf{\Phi}) \tag{2.114}$$

where

$$\mathbf{\Phi} \equiv \sum_i \frac{\partial \boldsymbol{f}}{\partial C_j}\dot{C}_j \tag{2.115}$$

is an identity, $\boldsymbol{f}(t, C_1, \ldots, C_6)$ and $\mathbf{g}(t, C_1, \ldots, C_6) \equiv \partial \boldsymbol{f}/\partial t$ being known functions. Now (2.114–2.115) make an incomplete system of six first-order equations for nine variables $(C_1, \ldots, C_6, \Phi_1, \ldots, \Phi_3)$. So one has to impose three arbitrary conditions on C, Φ, for example as

$$\boldsymbol{\Phi} = \boldsymbol{\Phi}(t, C_1, \ldots, C_6). \tag{2.116}$$

This will result in a closed system of six equations for six variables C_j:

$$\sum_i \frac{\partial \mathbf{g}}{\partial C_i} \dot{C}_i = \Delta \mathbf{F}\left(\boldsymbol{f}(t, C_1, \ldots, C_6), \mathbf{g}(t, C_1, \ldots, C_6) + \boldsymbol{\Phi}\right) - \dot{\boldsymbol{\Phi}} \tag{2.117}$$

$$\sum_i \frac{\partial \mathbf{f}}{\partial C_i} \frac{\mathrm{d}C_i}{\mathrm{d}t} = \boldsymbol{\Phi}, \tag{2.118}$$

$\boldsymbol{\Phi} = \boldsymbol{\Phi}(t, C_1, \ldots, C_6)$ now being some fixed function (gauge).[13] A trivial choice is $\boldsymbol{\Phi}(t, C_1, \ldots, C_6) = 0$, and this is what is normally taken by default. This choice is only one out of infinitely many, and often is not optimal. Under an arbitrary, non-zero, choice of the function $\boldsymbol{\Phi}(t, C_1, \ldots, C_6)$, the system (2.117–2.118) will have some different solution $C_j(t)$. To get the appropriate solution for the Cartesian components of the position and velocity, one will have to use formulae

$$\boldsymbol{r} = \boldsymbol{f}(t, C_1, \ldots, C_6), \tag{2.119}$$

$$\dot{\boldsymbol{r}} \equiv \mathbf{v} = \mathbf{g}(t, C_1, \ldots, C_6) + \boldsymbol{\Phi}(t, C_1, \ldots, C_6), \tag{2.120}$$

Appendix 2. Precession of the equator of date relative to the equator of epoch

The afore introduced vector $\boldsymbol{\mu}$ is the precession rate of the equator of date relative to the equator of epoch. Let the inertial axes (X, Y, Z) and the corresponding unit vectors $(\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}})$ be fixed in space so that X and Y belong to the equator of epoch. A rotation within the equator-of-epoch plane by longitude h_p, from axis X, will define the line of nodes, x. A rotation about this line by an inclination angle I_p will give us the planetary equator of date. The line of nodes x, along with axis y naturally chosen within the equator-of-date plane, and with axis z orthogonal to this plane, will constitute the precessing coordinate system, with the appropriate basis denoted by $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$.

In the inertial basis $(\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}})$, the direction to the North Pole of date is given by

$$\hat{\mathbf{z}} = \left(\sin I_p \sin h_p, -\sin I_p \cos h_p, \cos I_p\right)^T \tag{2.121}$$

while the total angular velocity reads:

$$\boldsymbol{\omega}_{\text{total}}^{(\text{inertial})} = \hat{\mathbf{z}}\,\Omega_z + \boldsymbol{\mu}^{(\text{inertial})}, \tag{2.122}$$

[13] Generally, $\boldsymbol{\Phi}$ may depend also upon the variables' time derivatives of all orders: $\boldsymbol{\Phi}(t; C_i, \dot{C}_i, \ddot{C}_i, \ldots)$. This will give birth to higher time derivatives of C in subsequent developments and will require additional initial conditions, beyond those on $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$, to be fixed to close the system. So it is practical to accept (2.116).

the first term denoting the rotation about the precessing axis $\hat{\mathbf{z}}$, and the second term being the precession rate of $\hat{\mathbf{z}}$ relative to the inertial frame $(\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}})$. This precession rate is given by

$$\boldsymbol{\mu}^{(\text{inertial})} = \left(\dot{I}_p \cos h_p, \dot{I}_p \sin h_p, \dot{h}_p\right)^T, \tag{2.123}$$

because this expression satisfies $\boldsymbol{\mu}^{(\text{inertial})} \times \hat{\mathbf{z}} = \dot{\hat{\mathbf{z}}}$.

In a frame co-precessing with the equator of date, the precession rate will be represented by vector

$$\boldsymbol{\mu} = \hat{\mathbf{R}}_{i \to p}\, \boldsymbol{\mu}^{(\text{inertial})}, \tag{2.124}$$

where the matrix of rotation from the equator of epoch to that of date (i.e., from the inertial frame to the precessing one) is given by

$$\hat{\mathbf{R}}_{i \to p} = \begin{bmatrix} \cos h_p & \sin h_p & 0 \\ -\cos I_p \sin h_p & \cos I_p \sin h_p & \sin I_p \\ \sin I_p \sin h_p & -\sin I_p \sin h_p & \cos I_p \end{bmatrix}$$

From here we get the components of the precession rate, as seen in the co-precessing coordinate frame (x, y, z):

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T = \left(\dot{I}_p, \dot{h}_p \sin I_p, \dot{h}_p \cos I_p\right)^T. \tag{2.125}$$

References

1. Efroimsky, Michael, and Peter Goldreich. (2003). Gauge symmetry of the N-body problem in the Hamilton–Jacobi approach. *Journal of Mathematical Physics*, **44**, pp. 5958–5977 astro-ph/0305344.
2. Efroimsky, Michael, and Peter Goldreich (2004). Gauge freedom in the N-body problem of celestial mechanics. *Astronomy & Astrophysics*, **415**, pp. 1187–1199 astro-ph/0307130.
3. Euler, L. (1748). *Recherches sur la question des inegalites du mouvement de Saturne et de Jupiter, sujet propose pour le prix de l'annee*. Berlin. Modern edition: L. Euler *Opera mechanica et astronomica*. Birkhauser-Verlag, Switzerland, 1999.
4. Euler, L. (1753). *Theoria motus Lunae exhibens omnes ejus inaequalitates etc*. Impensis Academiae Imperialis Scientarum Petropolitanae. St. Petersburg, Russia 1753. Modern edition: L. Euler *Opera mechanica et astronomica*. Birkhauser-Verlag, Switzerland 1999.
5. Lagrange, J.-L. (1778). *Sur le Problème de la détermination des orbites des comètes d'après trois observations, 1-er et 2-ième mémoires*. Nouveaux Mémoires de l'Académie de Berlin, 1778. Later edition: *Œuvres de Lagrange*. **IV**, Gauthier-Villars, Paris 1869.
6. Lagrange, J.-L. (1783). *Sur le Problème de la détermination des orbites des comètes d'après trois observations, 3-ième mémoire*. Ibidem, 1783. Later edition: *Œuvres de Lagrange*. **IV**, Gauthier-Villars, Paris 1869.
7. Lagrange, J.-L. (1808). Sur la théorie des variations des éléments des planètes et en particulier des variations des grands axes de leurs orbites. Lu, le 22 août 1808 à l'Institut de France. Later edition: *Œuvres de Lagrange*. **VI**, pp. 713–768, Gauthier-Villars, Paris 1877.
8. Lagrange, J.-L. (1809). Sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique. Lu, le 13 mars 1809 à l'Institut de France. Later edition: *Œuvres de Lagrange*. **VI**, pp. 771–805, Gauthier-Villars, Paris 1877.

9. Lagrange, J.-L. (1810). Second mémoire sur la théorie générale de la variation des constantes arbitraires dans tous les problèmes de la mécanique. Lu, le 19 février 1810 à l'Institut de France. Later edition: *Œuvres de Lagrange*. **VI**, pp. 809–816, Gauthier-Villars, Paris 1877.
10. Newman, W., and M. Efroimsky. (2003). The Method of Variation of Constants and Multiple Time Scales in Orbital Mechanics. *Chaos*, **13**, pp. 476–485.
11. Gurfil, P., and Klein, I. (2006). Mitigating the Integration Error in Numerical Simulations of Newtonian Systems. Submitted to *The International Journal for Numerical Methods in Engineering*.
12. Efroimsky, Michael (2002a). Equations for the orbital elements. Hidden symmetry. Preprint 1844 of the Institute of Mathematics and its Applications, University of Minnesota http://www.ima.umn.edu/preprints/ feb02/feb02.html.
13. Efroimsky, Michael (2002b). The implicit gauge symmetry emerging in the N-body problem of celestial mechanics. astro-ph/0212245.
14. Efroimsky, M. (2004). Long-term evolution of orbits about a precessing oblate planet. The case of uniform precession. astro-ph/0408168 (This preprint is a very extended version of the published paper Efroimsky (2005). It contains all technical calculations omitted in the said publication.)
15. Goldreich, P. (1965). Inclination of satellite orbits about an oblate precessing planet. *The Astronomical Journal*, **70**, pp. 5–9.
16. Brumberg, V. A., L. S. Evdokimova, and N. G. Kochina. (1971). Analytical methods for the orbits of artificial satellites of the moon. *Celestial Mechanics*, **3**, pp. 197–221.
17. Slabinski, V. (2003). Satellite orbit plane perturbations using an Efroimsky gauge velocity. Talk at the 34th Meeting of the AAS Division on Dynamical Astronomy, Cornell University, May 2003.
18. Gurfil, P. (2004). Analysis of J_2–perturbed Motion using Mean Non-Osculating Orbital Elements. *Celestial Mechanics & Dynamical Astronomy*, **90**, pp. 289–306.
19. Kinoshita, T. (1993). Motion of the Orbital Plane of a Satellite due to a Secular Change of the Obliquity of its Mother Planet. *Celestial Mechanics and Dynamical Astronomy*, **57**, pp. 359–368.
20. Efroimsky, M. (2005a). Long-term evolution of orbits about a precessing oblate planet. 1. The case of uniform precession. *Celestial Mechanics and Dynamical Astronomy*, **91**, pp. 75–108.
21. Lainey, V., Gurfil, P., and Efroimsky, M. (2005). Long-term evolution of orbits about a precessing oblate planet. 3. A semianalytical and a purely numerical approaches. *Celestial Mechanics and Dynamical Astronomy* (submitted).
22. Colombo, G. (1966). Cassini's second and third laws. *The Astronomical Journal*., **71**, pp. 891–896.
23. Efroimsky, M. (2005b). Long-term evolution of orbits about a precessing oblate planet. 2. The case of variable precession. *Celestial Mechanics & Dynamical Astronomy* (submitted).
24. Efroimsky, M. (2005c). The theory of canonical perturbations applied to attitude dynamics and to the Earth rotation. astro-ph/0506427.