

Solving Perturbed Kepler Problem by Quaternion Algebra

Jan Vrbik

Department of Mathematics, Brock University, St. Catharines, Canada

Email: jvrzik@brocku.ca

How to cite this paper: Vrbik, J. (2026) Solving Perturbed Kepler Problem by Quaternion Algebra. *Applied Mathematics*, 17, 337-353.

<https://doi.org/10.4236/am.2026.176021>

Received: March 31, 2026

Accepted: June 19, 2026

Published: June 22, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this article, we use quaternions together with Kustaanheimo-Stiefel transformation to solve the Kepler problem, first its unperturbed version and then when adding a small perturbing force. We assume that the perturbing force is autonomous, which enables us to derive a set of first-order differential equations for the corresponding orbital elements, while keeping them fully autonomous as well. This is achieved without compromising the resulting accuracy; there is no need for averaging out fast oscillations or involving a fast-angle variable. As a consequence, the equations can be solved to arbitrary accuracy by iterative and routine application of the new formulas.

Keywords

Quaternions, Kustaanheimo-Stiefel Transformation, Kepler Problem, Perturbation Theory, Orbital Elements

1. Quaternion Algebra

Quaternions are introduced as an extension of complex numbers where, instead of one imaginary unit i , there are three of them, denoted i , j and k ; we relate them to x , y and z directions of a three-dimensional space respectively. The square of each of these units equals to -1 , while any two of them anti-commute, e.g. $i \circ j = -j \circ i$ etc.; furthermore $i \circ j = k$, the small circle indicating quaternion multiplication (an *associative* operation) [1]. A quaternion is thus a quantity with four components, namely

$$\mathbb{A} := A + a_1 i + a_2 j + a_3 k = A + \mathbf{a} \quad (1)$$

where A is its *scalar* part and $\mathbf{a}$ its *vector* part (we use the blackboard type for full quaternions and boldface for vectors). Adding two quaternions is a component-wise operation, while multiplication is carried out (consistently with the pre-

vious rules) by

$$(A + \mathbf{a}) \circ (B + \mathbf{b}) = AB - \mathbf{a} \cdot \mathbf{b} + A\mathbf{b} + B\mathbf{a} + \mathbf{a} \times \mathbf{b} \quad (2)$$

where $\cdot$ and $\times$ denote the dot and vector product, respectively; quaternion multiplication is associative (as stated already) but *not* commutative. The proof of both statements is routine; it is based on the following two identities

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \quad (3)$$

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$$

Quaternion *conjugation* is defined (and denoted) by

$$\bar{\mathbb{A}} := A - \mathbf{a} \quad (4)$$

implying that $a := \sqrt{\mathbf{a} \circ \bar{\mathbf{a}}}$ yields the *length* of vector $\mathbf{a}$. Note that

$$\overline{\mathbb{A} \circ \mathbb{B}} = \bar{\mathbb{B}} \circ \bar{\mathbb{A}} \quad (5)$$

and that $\mathbb{A} \circ \bar{\mathbb{A}} = \bar{\mathbb{A}} \circ \mathbb{A}$ yields the sum of squares of $\mathbb{A}$'s four components (we may then drop the $\circ$ symbol and use $\mathbb{A}\bar{\mathbb{A}}$ or $\bar{\mathbb{A}}\mathbb{A}$ to emphasize that the result is a scalar and the order irrelevant). Using the conjugate, we then build the *inverse* of $\mathbb{A}$ by

$$\mathbb{A}^{-1} := \frac{\bar{\mathbb{A}}}{\mathbb{A}\bar{\mathbb{A}}} \quad (6)$$

so that $\mathbb{A} \circ \mathbb{A}^{-1} = \mathbb{A}^{-1} \circ \mathbb{A} = 1$.

Finally $\mathbb{A}^n$, where n is a positive integer, implies $\mathbb{A}$ multiplied by itself n times (calling the result the n^{th} power of $\mathbb{A}$); this means that we can now define a *function* of a quaternion, based on its Maclaurin expansion. The most useful of these is

$$\exp(\mathbb{A}) = \exp(A) \exp(\mathbf{a}) = \exp(A) \exp(a\hat{\mathbf{a}}) = \exp(A) (\cos a + \hat{\mathbf{a}} \sin a) \quad (7)$$

where $\hat{\mathbf{a}} := \frac{\mathbf{a}}{a}$ is the *unit direction* of $\mathbf{a}$; note that powers of $\hat{\mathbf{a}}$ follow the

$\hat{\mathbf{a}}, -1, -\hat{\mathbf{a}}, 1, \hat{\mathbf{a}}, \dots$ cycle, due to $\mathbf{a} \circ \mathbf{a} = -a^2$.

Rotation

Fixed rotation: Let $\mathbb{R}$ be a quaternion of unit magnitude, *i.e.* $\mathbb{R} := R + r\hat{\mathbf{r}}$ where $R^2 + r^2 = 1$ (equivalent to $\mathbb{R}\bar{\mathbb{R}} = 1$). This further implies that $\mathbb{R}$ can be uniquely written as

$$\mathbb{R} = \exp\left(\frac{\mathbf{w}}{2}\right) \quad (8)$$

where $\mathbf{w}$ is a vector whose direction is $\hat{\mathbf{r}}$ and whose magnitude meets

$\cos\left(\frac{w}{2}\right) = R$ and $\sin\left(\frac{w}{2}\right) = r$. When $\mathbf{x}$ represents a point in a three-dimensional (3D) space (using an *inertial* frame for its coordinates), it is easy to show that

$$\begin{aligned}\mathbf{x}_R &:= \mathbb{R} \circ \mathbf{x} \circ \bar{\mathbb{R}} = \exp\left(\frac{\mathbf{w}}{2}\right) \circ \mathbf{x} \circ \exp\left(-\frac{\mathbf{w}}{2}\right) \\ &= \mathbf{x}_{\parallel} + (\cos w + \hat{\mathbf{w}} \sin w) \circ \mathbf{x}_{\perp} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp} \cos w + \hat{\mathbf{w}} \times \mathbf{x}_{\perp} \sin w\end{aligned}\quad (9)$$

where $\mathbf{x}_R$ represents $\mathbf{x}$ rotated by angle w around the axis whose direction is $\hat{\mathbf{w}}$. Here $\mathbf{x}_{\parallel} := (\mathbf{x} \cdot \hat{\mathbf{w}}) \cdot \hat{\mathbf{w}}$ is the component of $\mathbf{x}$ parallel to $\hat{\mathbf{w}}$, while $\mathbf{x}_{\perp} := \mathbf{x} - (\mathbf{x} \cdot \hat{\mathbf{w}}) \cdot \hat{\mathbf{w}}$ is perpendicular to it. Note that $\mathbf{x}_{\parallel}$ commutes with $\mathbf{w}$ (and therefore with $\mathbb{R}$) while $\mathbf{x}_{\perp}$ anti-commutes (implying that $\mathbf{x}_{\perp} \circ \bar{\mathbb{R}} = \mathbb{R} \circ \mathbf{x}_{\perp}$).

When $\mathbf{x}(p)$ is a parametric description of a 3D object (such as an ellipse), the same operation similarly rotates the whole object, without changing its size or shape.

Uniform rotation: Changing $\mathbb{R}$ to

$$\mathbb{R} = \exp\left(\frac{\mathbf{w}}{2}s\right) \quad (10)$$

(where s is interpreted as time) makes the rotation time dependent; $\mathbf{w}$ then becomes the corresponding (constant) *angular velocity*.

Time-dependent rotation: Assuming that all components of $\mathbb{R}$ are now arbitrary functions of s (yet preserving the $\mathbb{R}\bar{\mathbb{R}} = 1$ property), we now find the corresponding *instantaneous* angular velocity at time s ; it follows from

$$\begin{aligned}\mathbf{x}'_R(s) &= \mathbb{R}' \circ \mathbf{x} \circ \bar{\mathbb{R}} + \mathbb{R} \circ \mathbf{x} \circ \bar{\mathbb{R}}' \\ &= \mathbb{R}' \circ \bar{\mathbb{R}} \circ \mathbf{x}_R(s) + \mathbf{x}_R(s) \circ \mathbb{R} \circ \bar{\mathbb{R}}' \\ &= 2(\mathbb{R}' \circ \bar{\mathbb{R}}) \times \mathbf{x}_R(s)\end{aligned}\quad (11)$$

since $\mathbb{R} \circ \bar{\mathbb{R}}' = -\mathbb{R}' \circ \bar{\mathbb{R}}$ (showing that $\mathbb{R}' \circ \bar{\mathbb{R}}$ is a *vector*) and since $\mathbf{a} \circ \mathbf{b} - \mathbf{b} \circ \mathbf{a} = 2\mathbf{a} \times \mathbf{b}$; this implies that $2\mathbb{R}' \circ \bar{\mathbb{R}}$ is the resulting velocity.

Kepler frame: Letting the original coordinate frame itself rotate with $\mathbb{R}$, and expressing the vector of the instantaneous angular velocity of the $\mathbb{R}$ rotation in this *new* (no longer inertial, but very convenient) *Kepler* frame, we get

$$\mathbf{Z} := 2\bar{\mathbb{R}} \circ (\mathbb{R}' \circ \bar{\mathbb{R}}) \circ \mathbb{R} = 2\bar{\mathbb{R}} \circ \mathbb{R}' \quad (12)$$

Euler angles: The usual way of parametrizing $\mathbb{R}$ is done [2] by

$$\mathbb{R} = \exp\left(\mathbf{k} \frac{\phi}{2}\right) \circ \exp\left(\mathbf{i} \frac{\theta}{2}\right) \circ \exp\left(\mathbf{k} \frac{\psi}{2}\right) \quad (13)$$

i.e. first rotating with respect to the z axis by angle ψ , then rotating with respect to the (original) x axis by angle θ , and finally rotating with respect to the z axis again by angle ϕ ; the three angles are referred to as *Euler angles* [3].

Using this form of $\mathbb{R}$ and letting each of the three angles be a function of s , the angular velocity (12), *i.e.* its Kepler-frame representation, becomes

$$\mathbf{Z} = \mathbf{i}(\theta' \cos \psi + \phi' \sin \psi \sin \theta) + \mathbf{j}(\phi' \cos \psi \sin \theta - \theta' \sin \psi) + \mathbf{k}(\psi' + \phi' \cos \theta) \quad (14)$$

We can then express each of the three time derivatives in terms of components of $\mathbf{Z}$ by solving the corresponding three equations, thus getting

$$\begin{aligned}\phi' &= \frac{Z_1 \sin \psi + Z_2 \cos \psi}{\sin \theta} \\ \theta' &= Z_1 \cos \psi - Z_2 \sin \psi \\ \psi' &= Z_3 - \phi' \cos \theta\end{aligned}\quad (15)$$

2. Kepler Problem

To find the motion of a satellite (of negligible mass) orbiting a primary of gravitational mass μ requires solving the following differential equation

$$\ddot{\mathbf{r}} + \mu \frac{\mathbf{r}}{r^3} = \varepsilon \mathbf{f} \quad (16)$$

where $\mathbf{r}$ is a vector with three components (x , y and z), each a function of time t , r is the length of $\mathbf{r}$, the double dot implies taking a second derivative of each of the $\mathbf{r}$ components with respect to t , and $\varepsilon \mathbf{f}$ is a *perturbing force* per unit mass of the satellite, assumed to be relatively small (*i.e.* $\varepsilon \ll \frac{\mu}{r^2}$) [3]; the

factor ε then appears in all quantities proportional to ε ; higher powers of ε are thus used to indicate the degree of smallness. When building a solution to (16), it will be important to distinguish between terms of the ε^0 type (solving the equation with no perturbing force), those proportional to ε , and finally terms proportional to second or higher power of ε (ultimately to be ignored, as explained shortly).

The article considers only *autonomous* perturbing forces, meaning that $\mathbf{f}$ is a function of $\mathbf{r}$ but *not* explicitly of time.

To utilize quaternion algebra for solving (16), we introduce a new *dependent* variable $\mathbb{U}$ (a quaternion) and a new *independent* variable s (a scalar) called *modified time* [4] and [5]. These are related to the old variables by

$$\mathbf{r} := \mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}} \quad (17a)$$

$$\frac{dt}{ds} := 2r \sqrt{\frac{a}{\mu}} \quad (17b)$$

where $a > 0$ is assumed (by subsequent proof) to be an *arbitrary* scalar function of s , and $r = \sqrt{-\mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}} \mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}}} = \mathbb{U} \bar{\mathbb{U}}$. We also define the following *scalar* function of s

$$\Gamma := \mathbb{U}' \circ \mathbf{i} \circ \bar{\mathbb{U}} - \mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}}' \quad (18)$$

where $'$ indicates differentiating with respect to s . Note that Γ is thus twice the scalar part of $\mathbb{U}' \circ \mathbf{i} \circ \bar{\mathbb{U}}$ (as an alternate definition).

One can then show that, using the new variables, (16) reads

$$\begin{aligned}2\mathbb{U}'' - \frac{2\mathbb{U}'\bar{\mathbb{U}}' - 4a}{r}\mathbb{U} - \frac{a'}{a}\left(\mathbb{U}' - \frac{\Gamma}{2r}\mathbb{U} \circ \mathbf{i}\right) + \frac{2\Gamma}{r}\mathbb{U}' \circ \mathbf{i} + \left(\frac{\Gamma}{r}\right)' \mathbb{U} \circ \mathbf{i} \\ = -4r \frac{a}{\mu} \varepsilon \mathbf{f} \circ \mathbb{U} \circ \mathbf{i}\end{aligned}\quad (19)$$

Proof:

$$\dot{\mathbf{r}} = \frac{2\mathbb{U}' \circ \mathbf{i} \circ \bar{\mathbb{U}} - \Gamma}{2r\sqrt{\frac{a}{\mu}}} = \frac{2\mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}}' + \Gamma}{2r\sqrt{\frac{a}{\mu}}} \quad (20)$$

due to (17b). Post-multiplying by $2\sqrt{\frac{a}{\mu}}\mathbb{U} \circ \mathbf{i}$ results in

$$2\sqrt{\frac{a}{\mu}}\dot{\mathbf{r}} \circ \mathbb{U} \circ \mathbf{i} = -2\mathbb{U}' - \frac{\Gamma}{r}\mathbb{U} \circ \mathbf{i} \quad (21)$$

Differentiating each side with respect to s yields

$$\begin{aligned} & \frac{a'}{a}\sqrt{\frac{a}{\mu}}\dot{\mathbf{r}} \circ \mathbb{U} \circ \mathbf{i} + 2\sqrt{\frac{a}{\mu}}\dot{\mathbf{r}} \circ \mathbb{U}' \circ \mathbf{i} + 4r\frac{a}{\mu}\left(\varepsilon\mathbf{f} - \mu\frac{\mathbf{r}}{r^3}\right) \circ \mathbb{U} \circ \mathbf{i} \\ &= -2\mathbb{U}'' - \left(\frac{\Gamma}{r}\right)' \mathbb{U} \circ \mathbf{i} - \frac{\Gamma}{r}\mathbb{U}' \circ \mathbf{i} \end{aligned} \quad (22)$$

which can be further simplified (note that $\mathbf{r} \circ \mathbb{U} \circ \mathbf{i} = \frac{\mathbf{r} \circ \mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}} \circ \mathbb{U}}{r} = r\mathbb{U}$) to get

$$\begin{aligned} & -\frac{a'}{a}\left(\mathbb{U}' + \frac{\Gamma}{2r}\mathbb{U} \circ \mathbf{i}\right) - \frac{2}{r}\bar{\mathbb{U}}'\mathbb{U}' \circ \mathbb{U} - \frac{\Gamma}{r}\mathbb{U}' \circ \mathbf{i} + 4r\frac{a}{\mu}\varepsilon\mathbf{f} \circ \mathbb{U} \circ \mathbf{i} + \frac{4a}{r}\mathbb{U} \\ &= -2\mathbb{U}'' - \left(\frac{\Gamma}{r}\right)' \mathbb{U} \circ \mathbf{i} - \frac{\Gamma}{r}\mathbb{U}' \circ \mathbf{i} \end{aligned} \quad (23)$$

which agrees with (19). ■

Note that *post*-multiplying $\mathbb{U}$ by $\exp(i\delta)$, where δ is any scalar function of s , does not change the resulting $\mathbf{r}$ and r ; this implies that $\mathbb{U} \circ \exp(i\delta)$ must solve (19) whenever $\mathbb{U}$ does (as can be explicitly verified). Nevertheless, the new solution does change the value of Γ to

$$\Gamma + \mathbb{U} \circ i\delta' \circ \mathbf{i} \circ \bar{\mathbb{U}} - \mathbb{U} \circ \mathbf{i} \circ (-i\delta') \circ \bar{\mathbb{U}} = \Gamma - 2r\delta' \quad (24)$$

This implies that we can always find a solution which meets $\Gamma = 0$ at all values of s by a proper choice of δ (referred to as *gauge*). Having done that, (19) is thus reduced to

$$2\mathbb{U}'' - \frac{2\mathbb{U}'\bar{\mathbb{U}}' - 4a}{r}\mathbb{U} - \frac{a'}{a}\mathbb{U}' = -4r\frac{a}{\mu}\varepsilon\mathbf{f} \circ \mathbb{U} \circ \mathbf{i} \quad (25)$$

while meeting $\Gamma = 0$. Since (25) and the last condition imply that Γ' is also identically equal to 0 (just post-multiply each term of (25) by $\mathbf{i} \circ \bar{\mathbb{U}}$ to show that $2\mathbb{U}'' \circ \mathbf{i} \circ \bar{\mathbb{U}}$ has zero scalar part), it is then sufficient to make Γ equal to 0 *initially*; solving (25) then assures that Γ remains constant, thus maintaining the $\Gamma = 0$ condition automatically at any future time s .

Unperturbed solution: When $\mathbf{f} = 0$, we choose a to be a positive *constant* and, assuming the solution to start with $\Gamma = 0$, the equation to solve is then

$$\mathbb{U}'' = \frac{\mathbb{U}'\bar{\mathbb{U}}' - 2a}{r}\mathbb{U} \quad (26)$$

where

$$\frac{\mathbb{U}'\bar{\mathbb{U}}' - 2a}{r} := E = \frac{2a}{\mu} \left(\frac{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}{2} - \frac{\mu}{r} \right) \quad (27)$$

is a constant of motion, proportional to the total (kinetic and gravitational) energy of the satellite [3].

Proof: Differentiating the LHS of (27) yields

$$\begin{aligned} & \frac{\mathbb{U}'' \circ \bar{\mathbb{U}}' + \mathbb{U}' \circ \bar{\mathbb{U}}''}{r} - \frac{E}{r} r' \\ &= \frac{E \mathbb{U} \circ \bar{\mathbb{U}}' + \mathbb{U}' \circ E \bar{\mathbb{U}}}{r} - \frac{E}{r} (\mathbb{U}' \circ \bar{\mathbb{U}} + \mathbb{U} \circ \bar{\mathbb{U}}') = 0 \end{aligned} \quad (28)$$

due to (26) and confirming that E' is identically zero. Furthermore

$$\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} = - \frac{\mathbb{U}' \circ \mathbf{i} \circ \bar{\mathbb{U}}}{r \sqrt{\frac{a}{\mu}}} \circ \frac{\mathbb{U} \circ \mathbf{i} \circ \bar{\mathbb{U}}'}{r \sqrt{\frac{a}{\mu}}} = \frac{\mu}{a} \frac{\mathbb{U}' \bar{\mathbb{U}}'}{r} \quad (29)$$

verifies the rest. ■

Assuming that E is negative (to make the satellite *orbit* the primary), we may now choose a so that $E = -1$. The equation to solve is then

$$\mathbb{U}'' + \mathbb{U} = \mathbb{O} \quad (30)$$

($\mathbb{O}$ denoting a zero quaternion) while

$$a = \frac{\mathbb{U}' \bar{\mathbb{U}}' + r}{2} \quad (31)$$

Since there is no *normal* (perpendicular to the plane defined by the initial $\mathbf{r}$ and $\dot{\mathbf{r}}$ directions) force, it is obvious that the motion, and the corresponding solution to (16), must be *planar*. It is thus sufficient to start with a solution that keeps $\mathbf{r}$ in the x - y plane; this solution then can be transformed (by applying an $\mathbb{R} = \exp\left(\mathbf{i} \frac{\theta}{2}\right) \circ \exp\left(\mathbf{k} \frac{\phi}{2}\right)$ fixed rotation to it) to any other specific attitude (the transformation preserves the $\Gamma = 0$ condition and the value of E).

To construct the most general solution to (30) [6], we use the following notation

$$\mathbf{q} := \exp(\mathbf{k}(s - \sigma)) \quad (32)$$

and

$$\mathbf{z} := \mathbf{q}^2 = \exp(2\mathbf{k}(s - \sigma)) \quad (33)$$

where σ is a constant. From now on, quantities of this type (*i.e.* having only scalar and $\mathbf{k}$ components) are called *complex* (with $\mathbf{k}$ representing the purely-imaginary unit) and can be multiplied using rules of complex algebra (no need for the $\circ$ operation, unless one of the factors is a quaternion); their complex nature is emphasized by using a different font.

The most general form of $\mathbb{U}$ that keeps (17a) in the x - y plane and meets (30) is then

$$U = \exp\left(\mathbf{k} \frac{\psi}{2}\right) (A\mathbf{q} + B\mathbf{q}^{-1}) \quad (34)$$

with A, B, σ and ψ being real constants; this implies that

$$r = A^2 + B^2 + AB(z + z^{-1}) \quad (35)$$

$$U' = \mathbf{k} \exp\left(\mathbf{k} \frac{\psi}{2}\right) (A\mathbf{q} - B\mathbf{q}^{-1}) \quad (36)$$

and

$$U' \bar{U}' = A^2 + B^2 - AB(z + z^{-1}) \quad (37)$$

further implying, due to (31), that

$$a = A^2 + B^2 \quad (38)$$

It is easy to see that $U' \circ \mathbf{i} \circ \bar{U}$ then has only $\mathbf{i}$ and $\mathbf{j}$ components, implying that $\Gamma = 0$ and confirming our original assumption.

The fully general solution is then

$$\mathbb{U} = \mathbb{R} \circ (A\mathbf{q} + B\mathbf{q}^{-1}) \quad (39)$$

where $\mathbb{R}$ is an arbitrary rotation. This implies that

$$\begin{aligned} \mathbf{r} &= \mathbb{R} \circ (A\mathbf{q} + B\mathbf{q}^{-1})^2 \circ \mathbf{i} \circ \bar{\mathbb{R}} \\ &= \mathbb{R} \circ (A^2 z + B^2 z^{-1} + 2AB) \circ \mathbf{i} \circ \bar{\mathbb{R}} \\ &= \mathbb{R} \circ \left(((A^2 + B^2) \cos(2s - 2\sigma) + 2AB) \mathbf{i} + (A^2 - B^2) \sin(2s - 2\sigma) \mathbf{j} \right) \circ \bar{\mathbb{R}} \end{aligned} \quad (40)$$

We already know that $A^2 + B^2 = a$; introducing $e := \frac{2AB}{A^2 + B^2}$, the same solution can be written as

$$\mathbf{r} = \mathbb{R} \circ a \left((\cos(2s - 2\sigma) + e) \mathbf{i} + \sqrt{1 - e^2} \sin(2s - 2\sigma) \mathbf{j} \right) \circ \bar{\mathbb{R}} \quad (41)$$

which is easily recognized as an equation of an ellipse with semi-major axis equal to a and eccentricity equal to e . The remaining parameters (called orbital elements) of the general solution are the three Euler angles of the ellipse's attitude and σ , the value of s at apocenter ($\mathbf{r}$ with the largest magnitude).

Finally, since

$$r = a(1 + e \cos(2s - 2\sigma)) \quad (42)$$

we get, based on (17b),

$$t = \frac{2a^{3/2}}{\sqrt{\mu}} (2s - 2\sigma + e \cos(2s - 2\sigma)) \quad (43)$$

relating real time t to $2s - 2\sigma$; the latter being a modified version (after subtracting π) of so called *eccentric anomaly*.

3. Perturbed Equation

When εf is nonzero, finding a solution becomes substantially more challeng-

ing; we can achieve it only to ε accuracy (when terms proportional to higher powers of ε are ignored, as done from now on). Nevertheless, once such first-order solution is found, we can use the same formulas to build a ε^2 -accurate (and higher) solution, by iterating.

To find ε -accurate solution [6], we must abandon the $\Gamma = 0$ gauge and return to solving (19) directly; to make the solution unique, a different gauge will offer itself in the process. We must also allow all six orbital elements of the unperturbed solution

$$\mathbb{U} = \mathbb{R} \circ \sqrt{\frac{a}{1+\beta^2}} (\mathbf{q} + \beta \mathbf{q}^{-1}) := \mathbb{R} \circ \mathbb{U}_0 \quad (44)$$

where $\beta := \frac{B}{A}$ (a more convenient parametrization) to be slowly varying (their *derivatives* proportional to ε) functions of s (this implies that their *second* derivatives are ε^2 -proportional and therefore ignored). The 0 subscript of $\mathbb{U}_0$ indicates that (i) the corresponding quantity is in Kepler's frame and (ii) it is evaluated to ε^0 (*i.e.* unperturbed-solution) accuracy. Note that now

$$\mathbf{q}' = k\mathbf{q}(1 - \varepsilon\sigma') \quad (45)$$

$$\mathbf{z}' = 2k\mathbf{z}(1 - \varepsilon\sigma')$$

(recall that ε multiplies all ε^1 -small quantities, while ε^2 -small and higher order terms are discarded).

Finally, $\mathbb{U}_0$ itself needs to be extended to (and replaced by)

$$\mathbb{U}_p := \mathbb{U}_0 + \varepsilon \mathbb{U}_D + \varepsilon \mathbb{U}_S \circ \mathbf{i} := \sqrt{\frac{a}{1+\beta^2}} \left(\mathbf{q} + \beta \mathbf{q}^{-1} + \mathbf{q} \varepsilon D(\mathbf{z}) + \mathbf{q} \frac{\varepsilon S(\mathbf{z}) + k\varepsilon b}{1+\beta z} \circ \mathbf{i} \right) \quad (46)$$

where $D(\mathbf{z})$ is a Laurent-series function of $\mathbf{z}$ [7], with missing $\mathcal{D}_{-1}z^{-1}$ and $\mathcal{D}_0z^0$ terms, *i.e.*

$$D(\mathbf{z}) := \cdots + \mathcal{D}_4 z^{-4} + \mathcal{D}_3 z^{-3} + \mathcal{D}_2 z^{-2} + \mathcal{D}_1 z + \mathcal{D}_2 z^2 + \mathcal{D}_3 z^3 + \mathcal{D}_4 z^4 + \cdots \quad (47)$$

where the $\mathcal{D}_j$ coefficients are (yet to be solved for) complex quantities.

Similarly, $S(\mathbf{z})$ is another such complex function which must furthermore meet the $S(\mathbf{z}) = -\overline{S(\mathbf{z})}$ condition (our new gauge), and is missing $\mathcal{S}_1 z$ (consequently $\mathcal{S}_{-1} z^{-1}$) and $\mathcal{S}_0$ terms. The latter one is actually needed to complete our solution, but is kept separate from S (we have denoted it $k b$, where b is real). Note that $\mathbb{U}_0$, $\mathbb{U}_D$ and $\mathbb{U}_S$ are thus complex quantities as well. The solution we hope to construct should then results in *unique* formulas for $\mathcal{D}_j$, $\mathcal{S}_j$, b and for the s derivatives of six orbital elements; we now proceed to find these by substituting the proposed solution (namely $\mathbb{R} \circ \mathbb{U}_p$) into (19), and matching its two sides.

Once done, (17a) would then convert the solution to $\mathbf{r}$ (as a function of s), while (17b) then relates s to real time t ; this is a routine final step of the procedure.

3.1. Further Simplification

We start by simplifying the solution's first two s -derivatives, getting

$$(\mathbb{R} \circ \mathbb{U}_p)' = \mathbb{R} \circ \mathbb{U}_p' + \mathbb{R}' \circ \mathbb{U}_p = \mathbb{R} \circ \left(\mathbb{U}_p' + \frac{\varepsilon \mathbf{Z}}{2} \circ \mathbb{U}_0 \right) \quad (48)$$

and

$$(\mathbb{R} \circ \mathbb{U}_p)'' = \mathbb{R} \circ (\mathbb{U}_p'' + \varepsilon \mathbf{Z} \circ \mathbb{U}_0') \quad (49)$$

We then simplify the resulting $\Gamma = \mathbb{U}_p' \circ \mathbf{i} \circ \overline{\mathbb{U}_p} - \mathbb{U}_p \circ \mathbf{i} \circ \overline{\mathbb{U}_p'}$, which turns out to be nonzero and ε -small. More specifically

$$\begin{aligned} \Gamma &= \mathbb{U}_p' \circ \mathbf{i} \circ \overline{\mathbb{U}_p} - \mathbb{U}_p \circ \mathbf{i} \circ \overline{\mathbb{U}_p'} + \varepsilon \frac{\mathbf{Z} \circ \mathbf{r}_0 + \mathbf{r}_0 \circ \mathbf{Z}}{2} \\ &= \varepsilon \left(\mathbb{U}_0' \overline{\mathbb{U}_s} - \mathbb{U}_0 \overline{\mathbb{U}_s'} - \mathbb{U}_s \overline{\mathbb{U}_0} + \mathbb{U}_s \overline{\mathbb{U}_0'} - \frac{Z_1 + kZ_2}{2} \overline{\mathbb{U}_0}^2 - \frac{Z_1 - kZ_2}{2} \mathbb{U}_0^2 \right) \end{aligned} \quad (50)$$

where $\mathbf{r}_0 := \mathbb{U}_0 \circ \mathbf{i} \circ \overline{\mathbb{U}_0}$, and a bar now implies *complex* conjugation. Note that contributions from the $\mathbb{U}_D$ part of the solution have cancelled out, the individual terms are complex, but the final Γ is real.

Proof:

$$\begin{aligned} \mathbb{U}_p' \circ \mathbf{i} \circ \overline{\mathbb{U}_p} &= (\mathbb{U}_0' + \varepsilon \mathbb{U}_D' + \varepsilon \mathbb{U}_s' \circ \mathbf{i}) \circ \mathbf{i} \circ (\overline{\mathbb{U}_0} + \varepsilon \overline{\mathbb{U}_D} - \mathbf{i} \circ \varepsilon \overline{\mathbb{U}_s}) \\ &= \mathbb{U}_0' \circ \mathbf{i} \circ \overline{\mathbb{U}_0} + \varepsilon \mathbb{U}_0' \mathbb{U}_D \circ \mathbf{i} + \varepsilon \mathbb{U}_D' \mathbb{U}_0 \circ \mathbf{i} + \varepsilon \mathbb{U}_0 \overline{\mathbb{U}_s} - \varepsilon \mathbb{U}_s' \overline{\mathbb{U}_0} \end{aligned} \quad (51)$$

Adding the corresponding complex conjugate removes the first three terms, as $\mathbb{U}_0' \circ \mathbf{i} \circ \overline{\mathbb{U}_0}$ contributes the unperturbed Γ (which equals to 0), and conjugating the next two term reverses their sign. The last two terms (plus their conjugates) then yield the first four terms of (50). Similarly

$$\begin{aligned} \mathbf{Z} \circ \mathbf{r}_0 &= ((Z_1 + kZ_2) \circ \mathbf{i} + kZ_3) \circ \mathbb{U}_0 \circ \mathbf{i} \circ \overline{\mathbb{U}_0} \\ &= -(Z_1 + kZ_2) \overline{\mathbb{U}_0}^2 + \mathbf{j} \circ Z_3 \overline{\mathbb{U}_0}^2 \end{aligned} \quad (52)$$

Adding its conjugate, namely

$$-(Z_1 - kZ_2) \mathbb{U}_0^2 - Z_3 \mathbb{U}_0^2 \circ \mathbf{j} = -(Z_1 - kZ_2) \mathbb{U}_0^2 - \mathbf{j} \circ Z_3 \mathbb{U}_0^2 \quad (53)$$

then verifies the remaining terms of (50). ■

We are now ready to substitute $\mathbb{R} \circ \mathbb{U}_p$ into (19); matching the equation's two sides then yields formulas for the s -derivatives of orbital elements and for the coefficients of $\mathbb{D}$ and $\mathbb{S}$. The resulting set of first-order differential equations for orbital elements (still truly autonomous, implying no need for averaging, and unaffected by one-orbit oscillations) can then be integrated, yielding invaluable insights into their long-range behaviour. On the other hand, $\mathbb{D}$ and $\mathbb{S}$ and b components of the solution provide parallel and perpendicular (to Kepler frame) distortions, respectively, of a single orbit. Since no approximation is used to build an ε -accurate solution, the same procedure can be used (iteratively) to construct ε^2 -accurate (and higher-order) solutions, thus achieving arbitrary accuracy. The main advantage over similar techniques such as Lie-Deprit transformation [8] and

Hamiltonian canonical perturbation theory [9] is in relative simplicity of our (yet to be derived) formulas. Admittedly, this derivation is rather involved (as seen shortly), but once found, the same formulas can be routinely applied to any autonomous perturbation and carried out to any order of ε accuracy. The only mathematical tool needed is already mentioned Laurent-series expansion of complex functions.

To find explicit formulas for all unknowns of the proposed solution, we first pre-multiply each term of (19) by $\overline{\mathbb{R}}$; this means that all quantities are transformed into the orbit's Kepler frame, correspondingly simplifying the $\mathbb{R} \circ \mathbb{U}_p$ solution and its two s derivatives (48) and (49), while the RHS becomes

$$-4r_0 \frac{a}{\mu} \varepsilon \mathbf{f}_0 \circ \mathbb{U}_0 \circ \mathbf{i} \quad (54)$$

where r_0 and $\mathbf{f}_0 := \overline{\mathbb{R}} \circ \mathbf{f} \circ \mathbb{R}$ are evaluated using the *unperturbed* solution $\mathbb{U}_0$. We now proceed to de-couple the resulting equation by converting it into two *complex* equations (easier to deal with, as multiplication becomes *commutative*).

3.2. Building a Solution

To derive first of these equations, we start with the LHS of the Kepler-frame version of (19), namely (expressed to the ε^1 accuracy, and utilizing the fact that Γ , Γ' and a' are ε small)

$$\begin{aligned} & 2\left(\mathbb{U}_p'' + \varepsilon \mathbf{Z} \circ \mathbb{U}_0'\right) - \frac{2\left(\mathbb{U}_p' + \frac{\varepsilon \mathbf{Z}}{2} \circ \mathbb{U}_0\right) \circ \left(\overline{\mathbb{U}}_p' - \overline{\mathbb{U}}_0 \circ \frac{\varepsilon \mathbf{Z}}{2}\right) - 4a}{r} \mathbb{U}_p \\ & - \frac{\varepsilon a'}{a} \mathbb{U}_0' + \frac{2\varepsilon \Gamma}{r_0} \mathbb{U}_0' \circ \mathbf{i} + \left(\frac{\varepsilon \Gamma}{r_0}\right)' \mathbb{U}_0 \circ \mathbf{i} \end{aligned} \quad (55)$$

Post-multiplying by $\overline{\mathbb{U}}_p$ and keeping complex components of each term only, we get (to the same ε^1 accuracy)

$$2\left(\mathbb{U}_1'' + \mathbf{k} \varepsilon \mathbf{Z}_3 \mathbb{U}_0'\right) \overline{\mathbb{U}}_1 - 2\left(\mathbb{U}_1' + \mathbf{k} \frac{\varepsilon \mathbf{Z}_3}{2} \mathbb{U}_0\right) \left(\overline{\mathbb{U}}_1' - \mathbf{k} \frac{\varepsilon \mathbf{Z}_3}{2} \overline{\mathbb{U}}_0\right) + 4a - \frac{\varepsilon a'}{a} \mathbb{U}_0' \overline{\mathbb{U}}_0 \quad (56)$$

(where $\mathbb{U}_1 := \mathbb{U}_0 + \varepsilon \mathbb{U}_D$), while the RHS becomes

$$-4r_0 \frac{a}{\mu} (\varepsilon \mathbf{f}_0 \circ \mathbf{r}_0)_{\text{cx}} := -4\varepsilon r_0 (1 + \beta z) Q(z) \quad (57)$$

(the cx subscript implies keeping only scalar and $\mathbf{k}$ components), and $Q(z)$ is correspondingly defined Laurent series. The Mathematica program of **Figure 1** completes the task of converting the LHS (short of the $\frac{a}{1+\beta^2}$ factor) into a function of $D(z)$ and of s derivatives of four orbital elements (highlighted); note that ε^0 terms have cancelled out, as expected. The code is reasonably self-explanatory, even to people not familiar with Mathematica.

Explicit formulas for the corresponding unknowns of the perturbed solution

are then found (see **Figure 2**) by matching coefficients of all powers of z in (57). The program first solves for the $\mathcal{D}_n$ coefficient of the $D(z)$ expansion while ignoring the a' , β' , σ' , Z_3 terms (contributing to z^{-1} , z^0 and z^1 powers only). This requires *simultaneously* solving for $\mathcal{D}_n$, $\mathcal{D}_{-n}$ and their complex conjugates, while considering them to be algebraically independent (they must and do turn out to be consistent with each other). To build the first of these four equations, we collect (the first line of code) coefficients of z^n which $D(z)$ contributes to the LHS of **Figure 1**, e.g. $\overline{D(z)}$ contributes $\mathcal{D}_{-n}$, $2z^2(1+\beta z)D''(z)$ contributes $2n(n-1)\mathcal{D}_n + 2\beta(n-1)(n-2)\mathcal{D}_{n-1}$, etc., and subtract the coefficient of z^n after expanding the RHS of (57); the second equation is just the $n \rightarrow -n$ counterpart of the first one, further including complex conjugates of both. The four equations are then expanded in powers of β and solved in an iterative manner; a careful choice of the expression multiplying $Q(z)$ on the RHS of (57) has made the iterative solution terminate upon reaching β^2 -proportional terms (see the first part of the program, ending with a highlighted $\mathcal{D}_n$ solution). Note that the formula does not solve for $\mathcal{D}_{-1}$ and $\mathcal{D}_0$, implying that

$$D(z) := \sum_{n=-\infty}^{-2} \mathcal{D}_n z^n + \sum_{n=1}^{\infty} \mathcal{D}_n z^n \quad (58)$$

```
cc[A_] := A /. {Complex[x_, y_] -> Complex[x, -y],
  D[z[s]] -> Dbar[z[s]], D'[z[s]] -> -z[s]^2 Dbar'[z[s]],
  Dbar[z[s]] -> D[z[s]], Dbar'[z[s]] -> -z[s]^2 D'[z[s]],
  q[s] -> q[s]^-1, z[s] -> z[s]^-1};
cIn[x_] := Expand[x] /. e^n_ /; n > 1 -> 0 // Simplify;
q'[s] := 1 (1 - e sigma') q[s]; z'[s] := 2 1 (1 - e sigma') z[s];
Unprotect[Power]; q[s]^2 := z[s]; q[s]^-2 := z[s]^-1;
Uo = Sqrt[a[e s] / (1 + beta[e s]^2)] (q[s] + beta[e s] / q[s]);
U = Uo + e Sqrt[a / (1 + beta^2)] D[z[s]] x q[s];
U' = D[U, s] + 1 e Z3 Uo / 2 // cIn;
U'' = D[U', s] + 1 e Z3 D[Uo, s] / 2 // cIn;
(1 + beta^2) / a x (2 U'' cc[U] - 2 U' cc[U'] +
  4 a - e a' / a D[Uo, s] x cc[Uo]) // cIn;
Collect[% /. e -> 1 /. {x_[s] -> x}, {a', beta', sigma', Z3,
  Dbar[z], D'[z], Dbar'[z], D''[z]}, Simplify]
```

$$\left(\left(-4 + \frac{2\beta}{z} - 2z\beta + 4\beta^2 \right) Z_3 - 4D[z] - \frac{1(-\beta + z^2\beta + z(-1 + \beta^2))a'}{az} - \frac{21(1 + 4z\beta + \beta^2 + z^2(1 + \beta^2))\beta'}{z(1 + \beta^2)} + 8(1 + \beta^2)\sigma' - 4Dbar[z] - 4z(5 + 3z\beta)D'[z] + 4(z - \beta)Dbar'[z] - 8z^2(1 + z\beta)D''[z] \right)$$

Figure 1. (56) further simplified.

```

eq = (1 + 5 n + 2 n (n - 1)) Dn + (3 β (n - 1) + 2 β (n - 1) (n - 2)) Dn-1 +
      (1 - n) D-n + β (n + 1) D-n-1 - (1 + 2 β2) Qn - β Qn+1 - β (2 + β2) Qn-1 - β2 Qn-2;
Dn := d0[n] + β d1[n] + β2 d2[n]; D-n := d0[n] + β d1[n] + β2 d2[n];
cc[x_] :=
  x /. {d → D, D → d, Q → Q, Q → Q, D → D, D → D, z → 1/z, Complex[b_, c_] → Complex[b, -c]};
Do[cf = Coefficient[eq, β, i] == 0;
  sol = Solve[{cf, cf /. n → -n, cc[cf], cc[cf] /. n → -n}, {di[n], di[n], di[-n], di[-n]}];
  s = {sol[[1, 1, 2]], sol[[1, 2, 2]]} // Simplify;
  di[m_] = s[[1]] /. n → m; di[m_] = s[[2]] /. n → m, {i, 0, 2}]; Collect[Dn, β, Simplify]


$$\frac{\beta^2 ((3 + 2n) Q_n - Q_{-2-n})}{4n(1+n)^2} + \frac{\beta ((1 + 3n + 2n^2) Q_{-1+n} + n(1 + 2n) Q_{1+n} - Q_{-1-n})}{4n^2(1+n)^2} + \frac{(-1 + 2n) Q_n + Q_{-n}}{4n^2(1+n)}$$


D[z_] = d1 z + d-2 z-2; D[z_] = d1 / z + d-2 z2;
aux = -4 D[z] - 4 D[z] - 4 z (5 + 3 z β) D'[z] + 4 (z - β) D'[z] -
      8 z2 (1 + z β) D''[z] + 4 (1 + β z)2 (1 + β / z) Sum[Qi zi, {i, -3, 2}];
aux = aux + (-4 + 2 β / z - 2 z β + 4 β2) Z3 + i (β / z - z β + 1 - β2) a' / a -
      2 i (1 / z + 4 β + β2 / z + z + z β2) β' / (1 + β2) + 8 × (1 + β2) σ';
aux = Coefficient[aux, z, {-1, 0, 1}];
sol = Solve[Thread[aux + cc[aux] == 0], {Z3, σ'}];
Collect[sol[[1]] /. Qi → 2 Re[Qi] - Qi, Re[_], Simplify]


$$\left\{ \begin{aligned} Z_3 &\rightarrow \frac{1}{2} \times (-1 - \beta^2) \operatorname{Re}[Q_{-2}] + \frac{(-3 + \beta^2) \operatorname{Re}[Q_{-1}]}{2\beta} + \frac{1}{2} \times (-1 + 3\beta^2) \operatorname{Re}[Q_0] + \frac{(1 + \beta^2) \operatorname{Re}[Q_1]}{2\beta}, \\ \sigma' &\rightarrow \frac{(-1 - 2\beta^2 + \beta^4) \operatorname{Re}[Q_{-2}]}{4 \times (1 + \beta^2)} - \frac{3 \times (1 + \beta^4) \operatorname{Re}[Q_{-1}]}{4(\beta + \beta^3)} - \frac{3 \times (1 + \beta^4) \operatorname{Re}[Q_0]}{4 \times (1 + \beta^2)} - \frac{(-1 + 2\beta^2 + \beta^4) \operatorname{Re}[Q_1]}{4(\beta + \beta^3)} \end{aligned} \right\}$$


sol = Solve[Thread[aux - cc[aux] == 0], {a', β'}];
Collect[sol[[1]] /. Qi → Qi - 2 i Im[Qi], Im[_], Simplify]


$$\left\{ \begin{aligned} a' &\rightarrow 4 a \beta \operatorname{Im}[Q_{-1}] - 4 a \operatorname{Im}[Q_0], \\ \beta' &\rightarrow \frac{1}{2} (\beta + \beta^3) \operatorname{Im}[Q_{-2}] + \frac{3}{2} \times (1 + \beta^2) \operatorname{Im}[Q_{-1}] + \frac{3}{2} (\beta + \beta^3) \operatorname{Im}[Q_0] + \frac{1}{2} \times (1 + \beta^2) \operatorname{Im}[Q_1] \end{aligned} \right\}$$


```

Figure 2. Solving (56)=(57).

This makes $D(z)$ match coefficients of all powers of z , with the *exception* of z^{-1} , z^0 and z^1 ; these are then used to solve for a' , β' , σ' and Z_3 ; this is further simplified by solving, separately (yet another helpful de-coupling) the real (yielding Z_3 and σ') and purely imaginary (yielding a' and β') parts of the corresponding three equations. Note that in both cases the equations are linearly dependent, thus allowing for a unique solution of each pair of unknowns (highlighted at the end of Figure 3).

To find $S(z)$, Z_1 , Z_2 and b , we now keep only the i and j terms of (55), thus getting

$$2U_S'' \circ i + \varepsilon(Z_1 i + Z_2 j) \circ U_0' + 2U_S \circ i + \frac{2\varepsilon\Gamma}{r_0} U_0' \circ i + \left(\frac{\varepsilon\Gamma}{r_0}\right)' U_0 \circ i$$

Post-multiplying by $i \circ \overline{U_0}$ results in

```

cc[A_] := A /. {Complex[x_, y_] -> Complex[x, -y],
  \[Epsilon]''[z[s]] -> -z[s]^4 \[Epsilon]''[z[s]] - 2 z[s]^3 \[Epsilon]'[z[s]], \[Epsilon]'[z[s]] -> \[Epsilon]'[z[s]] z[s]^2,
  \[Epsilon][z[s]] -> -\[Epsilon][z[s]], q[s] -> q[s]^-1, z[s] -> z[s]^-1};
cIn[x_] := (x // Expand) /. {e^n_/: n>1 -> 0} // Simplify;
q'[s] := i (1 - e \[Epsilon]'[s]) q[s]; z'[s] := 2 i (1 - e \[Epsilon]'[s]) z[s];
Unprotect[Power]; q[s]^2 := z[s]; q[s]^-2 := z[s]^-1; q[s]^4 := z[s]^2;
Uo = (q[s] + \[Beta][s] / q[s]); Us = e (\[Epsilon][z[s]] + i b) \times q[s] / (1 + \[Beta] z[s]);
r = Uo cc[Uo] // cIn;
\Gamma = D[Uo, s] \times cc[Us] - cc[Uo] \times D[Us, s]; \Gamma = \Gamma + cc[\Gamma];
\Gamma = \Gamma - e (Z1 + I Z2) cc[Uo]^2 / 2 - e (Z1 - I Z2) Uo^2 / 2;
(2 (D[Us, {s, 2}] + Us) \times cc[Uo] + 2 e (Z1 + I Z2) cc[D[Uo, s] Uo] +
  2 \Gamma / r D[Uo, s] \times cc[Uo] + D[\Gamma / r, s] r) // cIn;
Collect[% /. e -> 1 /. {x_[s] -> x}, {Z1, Z2, \[Epsilon][z], \[Epsilon]'[z], b}, Simplify]

```

$$\frac{8 i b z (1 + \beta^2)}{(z + \beta) (1 + z \beta)} + \frac{i (-1 + \beta^2) (\beta + 6 z^2 \beta + z^4 \beta + 2 z (1 + \beta^2) + 2 z^3 (1 + \beta^2)) Z_1}{z (z + \beta) (1 + z \beta)} -$$

$$\frac{(-1 + z^2) (\beta + \beta^3 + z^2 (\beta + \beta^3) + 2 z (1 + \beta^4)) Z_2}{z (z + \beta) (1 + z \beta)} +$$

$$\frac{8 z (1 + \beta^2) \[Epsilon][z]}{(z + \beta) (1 + z \beta)} - \frac{8 z (z + 2 \beta + z \beta^2) \[Epsilon]'[z]}{(z + \beta) (1 + z \beta)} - 8 z^2 \[Epsilon]''[z]$$

Figure 3. LHS of (59).

$$\begin{aligned}
 & -2\varepsilon(U_s'' + U_s)\overline{U_0} - 2\varepsilon(Z_1 + kZ_2)\overline{U_0'}U_0 - \frac{2\varepsilon\Gamma}{r_0}U_0'\overline{U_0} - \left(\frac{\varepsilon\Gamma}{r_0}\right)'r_0 \\
 & = 4kr_0^2 \frac{a}{\mu}(\varepsilon f_0)_3 := 4\varepsilon r_0 W(z)
 \end{aligned} \tag{59}$$

where $\overline{W(z)} = -W(z)$ (note that each term of the equation has this property).

We then proceed to simplify the LHS (short of a $-\frac{a}{1+\beta^2}$ factor) of (59) by

Mathematica program of Figure 4.

And then to find S_n (now much easier, as only one equation needs to be solved), followed by deriving formulas for Z_1 , Z_2 and b (by matching coefficients of z^0 and z^1); this is done in Figure 5.

We now have a complete set of formulas needed to find (based on the perturbing force) all ingredients of our solution. Using this (ε -accurate) solution, we can then evaluate both sides of (19) to the ε^2 accuracy and use the same formulas to construct ε^2 -accurate solution, and so on. Post-multiplying the final solution by $\exp(i\delta \cdot s)$, where δ is a properly chosen *constant*, we can then impose the $\Gamma = 0$ gauge, if desired.

3.3. Oblateness Example

To illustrate how to apply these formulas, we consider the perturbing force experienced by an artificial satellites due to Earth's oblateness, given by (using Kepler's frame).


```

eq = 2 × (1 + β²) × (1 - n²) Sn - 2 β (n - 1) (n - 2) Sn-1 - 2 β (n + 2) (n + 1) Sn+1 - β² (Wn+2 + Wn-2) -
      2 β (1 + β²) (Wn-1 + Wn+1) - (1 + 4 β² + β⁴) Wn; Sn := s0[n] + β s1[n] + β² s2[n];
Do[cf = Coefficient[eq, β, i] == 0;
  sol = Solve[cf, si[n]];
  r = sol[[1, 1]] // Simplify;
  si[m_] = r[[2]] /. n → m, {i, 0, 2}]; Sn // Simplify

```

$$-\frac{(1+n)\beta W_{-1+n} + n(1+\beta^2)W_n + (-1+n)\beta W_{1+n}}{2n(-1+n^2)}$$

```

cc[x_] := x /. {W → W̄, W̄ → W, z → 1/z, Complex[b_, c_] → Complex[b, -c], W0 → -W0}
ϖ[z_] = Sum[Si zi - cc[Si] z-i, {i, 2, 4}];
aux = 8 i b (1 + β²) + (i (-1 + β²) (β + 6 z² β + z⁴ β + 2 z (1 + β²) + 2 z³ (1 + β²)) Z1 / z²
      - ((-1 + z²) (β + β³ + z² (β + β³) + 2 z (1 + β⁴)) Z2) / z²;
aux = aux + 8 × (1 + β²) ϖ[z] - 8 (z + 2 β + z β²) ϖ'[z] - 8 z (z + β) (1 + z β) ϖ''[z];
aux = aux - 4 × (1 + β z)² × (1 + β / z)² (Sum[Wi zi - W̄i z-i, {i, 4}] + W0);
aux = Collect[aux // Expand, z, Simplify]; eq = Coefficient[aux, z, {0, 1, 2}];
(Solve[Thread[eq == 0], {Z1, Z2, b}][[1]] /.
  {W0 → i Im[W0], -W1 - W̄1 → -2 Re[W1], W̄i → Wi - 2 i Im[Wi]} // Simplify

```

$$\left\{ Z_1 \rightarrow \frac{4\beta \operatorname{Im}[W_0] + 2 \times (1 + \beta^2) \operatorname{Im}[W_1]}{-1 + \beta^2}, Z_2 \rightarrow -2 \operatorname{Re}[W_1], \right.$$

$$\left. b \rightarrow \frac{(-1 + \beta^2)^2 \operatorname{Im}[W_0] + \beta ((1 + \beta^2) \operatorname{Im}[W_1] + 2\beta \operatorname{Im}[W_2])}{2 \times (1 + \beta^2)} \right\}$$

Figure 4. Solving (59).

```

cc[x_] := x /. {z → 1/z, Complex[b_, c_] → Complex[b, -c]}; Re[x_] := (x + cc[x]) / 2;
Im[x_] := (x - cc[x]) / 2 / i; r = a / (1 + β²) (1 + β / z) × (1 + β z);
r = a / (1 + β²) (z + β)² / z; u = Sin[θ] (Sin[ψ] + i Cos[ψ]); ur = cc[u] r;
Q = e a R² (3 / 2 r² - 9 / 2 Re[ur]² - 3 i Re[ur] × Im[ur]) / r⁵ / (1 + β z);
Qi = Residue[Q z-i-1, {z, -β}] // Simplify; a' → 4 a β Im[Q-1] - 4 a Im[Q0] // Simplify
β' → 1 / 2 (β + β³) Im[Q-2] + 3 / 2 × (1 + β²) Im[Q-1] + 3 / 2 (β + β³) Im[Q0] + 1 / 2 × (1 + β²) Im[Q1] // Simplify
Z3 → 1 / 2 × (-1 - β²) Re[Q-2] + (-3 + β²) Re[Q-1] / 2 β + 1 / 2 × (-1 + 3 β²) Re[Q0] + (1 + β²) Re[Q1] / 2 β //
Simplify
σ' → (-1 - 2 β² + β⁴) Re[Q-2] / 4 × (1 + β²) - 3 × (1 + β⁴) Re[Q-1] / 4 (β + β³) -
      3 × (1 + β⁴) Re[Q0] / 4 × (1 + β²) - (-1 + 2 β² + β⁴) Re[Q1] / 4 (β + β³) // Simplify

```

$$a' \rightarrow 0$$

$$\beta' \rightarrow 0$$

$$Z_3 \rightarrow \frac{3 R^2 (1 + \beta^2)^4 e (1 + 3 \cos[2 \theta])}{4 a^2 (-1 + \beta^2)^4}$$

$$\sigma' \rightarrow -\frac{3 R^2 (1 + \beta^2)^3 e (1 + 3 \cos[2 \theta])}{16 a^2 (-1 + \beta^2)^3}$$

Figure 5. Orbital-element derivatives under oblateness perturbations (Part 1).

$$f_0 := -\frac{\varepsilon \mu R^2}{r_0^5} \left(\frac{3}{2} r_0 - \frac{15 (r_0 \cdot u_0)^2}{2 r_0^2} r_0 + 3 (r_0 \cdot u_0) u_0 \right) \quad (60)$$

where ε is the second zonal-harmonic coefficient, R is the Earth's equatorial radius, and u (equal to k , by our choice of coordinates) is the unit direction of the Earth's axis. To convert u to Kepler's frame, we post-multiply it by $\mathbb{R}$ of (13) and pre-multiply by $\overline{\mathbb{R}}$, thus getting

$$u_0 = i \sin \theta \sin \psi + j \sin \theta \cos \psi + k \cos \theta \quad (61)$$

The key quantities needed to build the corresponding perturbed solution are

$$Q(z) = \frac{\varepsilon a R^2}{r_0^5 (1 + \beta z)} \left(\frac{3}{2} r_0^2 - \frac{9}{2} (r_0 \cdot u_0)^2 - 3k (r_0 \cdot u_0) (u_0 \times r_0)_3 \right) \quad (62)$$

and

$$W(z) = -\frac{\varepsilon a R^2}{r_0^6} 3k (r_0 \cdot u_0) \cos \theta \quad (63)$$

Converting these into derivatives of the orbital elements is done by Mathematical programs of **Figure 5** and **Figure 6**; while both a' and β' turn out to be equal to zero, further application of (15) results in the following classical formulas

$$\phi' = -\frac{2\varepsilon (1 + \beta^2)^4}{a^2 (1 - \beta^2)^4} \cos \theta \quad (64)$$

$$\theta' = 0$$

$$\psi' = \frac{\varepsilon (1 + \beta^2)^4}{a^2 (1 - \beta^2)^4} (5 \cos^2 \theta - 1)$$

```
cc[x_] := x /. {z -> 1/z, Complex[b_, c_] -> Complex[b, -c]}
Re[x_] := (x + cc[x])/2; Im[x_] := (x - cc[x])/2/i
r = a/(1 + beta^2) (1 + beta/z) x (1 + beta z); r = a/(1 + beta^2) (z + beta)^2/z;
u = Sin[theta] (Sin[psi] + i Cos[psi]); ur = cc[u] r;
W = epsilon a R^2 (3 i Re[ur] Cos[theta])/r^4;
W1_ = Residue[W z^-1, {z, -beta}] // Simplify;
Z1 -> (4 beta Im[W1_] + 2 x (1 + beta^2) Im[W1_]) / (-1 + beta^2) // ExpToTrig // Simplify
Z2 -> -2 Re[W1_] // ExpToTrig // Simplify
```

$$Z_1 \rightarrow -\frac{3 R^2 (1 + \beta^2)^4 \epsilon \sin[2 \theta] \sin[\psi]}{2 a^2 (-1 + \beta^2)^4}$$

$$Z_2 \rightarrow -\frac{3 R^2 (1 + \beta^2)^4 \epsilon \cos[\psi] \sin[2 \theta]}{2 a^2 (-1 + \beta^2)^4}$$

Figure 6. Orbital-element derivatives under oblateness perturbation (Part 2).

which translate into a slow, uniform precession of the orbital plane around the Earth's axis (with angular speed of ϕ') and a similar rotation of the orbit's perigee within the orbital plane with angular speed of ψ' , while keeping the inclination angle fixed. Note that the perigee rotation ceases when $\theta = \arcsin \frac{2}{\sqrt{5}}$, called the *critical angle*; the ε^2 -accurate solutions reveals that, instead of fully stopping, the rotation turns into a libration around the critical angle [10] and [11], but pursuing this goes beyond the scope of this article (consult the two references for further details). The first of these also derives expressions for resulting distortions of the Kepler's frame $\mathbf{r}$; we do not quote them here due to their complexity.

4. Conclusions

In this final section we want to mention a few potential extensions of the new technique (excluded from our basic presentation). Firstly, the technique is easily capable of dealing with more than one autonomous force at a time and (at the same time) reaching an arbitrarily accurate solution by iterating; this simply requires attaching a *different* small parameter ($\varepsilon_1, \varepsilon_2, \dots$) to each individual perturbing force and applying existing formulas to individual terms of the corresponding expansion. A more difficult task (even in a single-force situation) is to establish the largest value of ε which forces the iteration process to converge; this remains an open question. But luckily, in typical applications, perturbing forces are small enough to guarantee fast convergence.

The most obvious modification of our formulas is clearly needed when the perturbing force is no longer autonomous; this has been done in [12]. The solution can no longer avoid introducing fast (*i.e.* one orbit) oscillations into the resulting differential equations for orbital elements, thus making it more difficult to maintain high accuracy when exploring their long-term (millions of years) behavior. Nevertheless, in these cases, getting the correct qualitative picture is often the main goal, amply met by the extended formulas.

When all parameters of the perturbing force are numerical, it becomes beneficial (when constructing higher-order solution) to modify the iteration process not just to find the next-order solution, but also to *update* the existing one (till no changes are observed); this has been demonstrated in [13] and offers a way of properly dealing with *small divisors*.

The final challenge to solving a perturbed Kepler problem is due to so called *resonances*, happening when a ratio of the satellite's (used in a generic sense) orbital period to the period of a (cyclic) perturbing force becomes a simple fraction (such as 2:1, 3:2 etc.); the new technique is well equipped to elucidate main features of the corresponding solution, as corroborated by [14] and several related articles.

The new technique has thus a large range of applications (well beyond what we could cover in this article) due to its ability to resolve several potential issues adversely affecting many traditional methods of solution. Yet, further advancement

is still possible, but requires a solid background in the technique's fundamentals; it is hoped that our article has provided it.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Serre, J.P. (1973) A Course in Arithmetic, Graduate Texts in Mathematics. Springer.
- [2] Chelnokov, Y.N. (2022) Quaternion Methods and Models of Regular Celestial Mechanics and Astrodynamics. *Applied Mathematics and Mechanics*, **43**, 21-80. <https://doi.org/10.1007/s10483-021-2797-9>
- [3] Goldstein, H. (1980) Classical Mechanics. 2nd Edition, Addison-Wesley.
- [4] Kustaanheimo, P., Schinzel, A., Davenport, H. and Stiefel, E. (1965) Perturbation Theory of Kepler Motion Based on Spinor Regularization. *Journal für die reine und angewandte Mathematik*, **1965**, 204-219. <https://doi.org/10.1515/crll.1965.218.204>
- [5] Stiefel, E.L. and Scheifele, G. (1971) Linear and Regular Celestial Mechanics. Springer-Verlag.
- [6] Vrbik, J. (2023) New Methods of Celestial Mechanics. Bantam Science Publishers.
- [7] Rudin, W. (1987) Real and Complex Analysis. 3rd Edition, McGraw-Hill.
- [8] Cary, J.R. (1981) Lie Transform Perturbation Theory for Hamiltonian Systems. *Physics Reports*, **79**, 129-159. [https://doi.org/10.1016/0370-1573\(81\)90175-7](https://doi.org/10.1016/0370-1573(81)90175-7)
- [9] Boccaletti, D. and Pucacco, G. (1999) Theory of Orbits. Volume 2: Perturbative and Geometrical Methods. Springer-Verlag.
- [10] Vrbik, J. (1997) Oblateness Perturbations to Fourth Order. *Monthly Notices of the Royal Astronomical Society*, **291**, 65-70. <https://doi.org/10.1093/mnras/291.1.65>
- [11] Vrbik, J. (2009) Second Erratum: Oblateness Perturbations to Fourth Order. *Monthly Notices of the Royal Astronomical Society*, **399**, 1088. <https://doi.org/10.1111/j.1365-2966.2009.15412.x>
- [12] Vrbik, J. (1995) Perturbed Kepler Problem in Quaternionic Form. *Journal of Physics A: Mathematical and General*, **28**, 6245-6252. <https://doi.org/10.1088/0305-4470/28/21/027>
- [13] Vrbik, J. (2001) Quaternionic Processor. *Celestial Mechanics and Dynamical Astronomy*, **80**, 111-118. <https://doi.org/10.1023/a:1011979701759>
- [14] Vrbik, J. (1996) Resonance Formation of Kirkwood Gaps and Asteroid Clusters. *Journal of Physics A: Mathematical and General*, **29**, 3311-3316. <https://doi.org/10.1088/0305-4470/29/12/033>