rijksuniversiteit
groningen

faculteit Wiskunde en
Natuurwetenschappen

Algebraic approaches to regularizing the Kepler problem.

Bachelor Thesis Mathematics & Physics

July 2013

Student: G. S. P. Wiersema

First supervisor: Prof. dr. H Waalkens

Second supervisor: Prof. dr. ir. E van der Giessen

Abstract

This paper analyzes algebraic regularizations of the Kepler problem. The Kepler problem treats the motion of two bodies interacting through a central force that obeys the inverse square law. The paper first discusses the Kepler problem in detail. Then two methods of regularizing the Kepler problem are compared: Algebraic regularization of the Kepler problem can be achieved by using quaternions as shown in the paper by J. Walvogel (2007) [13], which is treated first. Then, algebraic regularization of the Kepler problem by using spinors is discussed, as shown in the paper by T. Bartsch (2003) [3]. The paper then continues to show that quaternions are isomorphic to spinors, which explains the many similarities between both regularizations. These similarities are studied in detail in the final part of this paper.

1 Introduction

The Kepler problem is named after Johannes Kepler, who formulated Kepler's law in order to describe the orbits of planets around the sun. However, the Kepler problem has since then turned up in many more situations and so the Kepler formulae resulted to be much more widely applicable then just in celestial mechanics.

The Kepler problem is a classical mechanical problem that treats the interaction of two bodies. The two bodies interact through a force $\vec{F}$, whose strength is inversely proportional to the square of the distance r between the two bodies. From the reference frame of one of the bodies, this body travels through the forcel field generated by the other body. This force field has a singularity at the origin. From a physical standpoint, we can see this from the fact that the origin of the force field is where the body generating the force field is located, so the force is not defined there. Mathematically, we can see from the inverse relationship between the strength of the force and the square of the distance to the origin, that at the origin (where the distance to the origin is zero) the strength of the force blows up. When solving for the orbit of the body in the force field near the origin, numerical solutions become unstable because the force that generates the oribit blows up there. Although the unpertubed Kepler problem can be solved exactly, when introducing pertubations such as an electric force field it can become useful to use numerical solutions. Thus, to facilitate numerical solutions near the singularity at the origin, we would like to regularize the Kepler problem.

This paper will analyze algebraic approaches to regularizing the Kepler problem. Two papers on reguralizing the Kepler problem will be analyzed, discussed and finally compared. The first section of this paper will define the Kepler problem and discuss its background, applications and solutions. The second section of this paper will concern algebraic regularization of the Kepler problem using quaternions. This section will begin with a summary of quaternion algebra. Subsequently, we will use the quaternion algebra to explain the algebraic regularization of the Kepler problem using quaternions, according to the paper by J. Walvogel (2007)[13]. The third section will treat the regularization of the Kepler problem by using spinors. This section will begin by explaining the use of geometric algebra and spinors as a preliminary to the part that follows; explaining algebraic regularization using spinors, as shown in the paper by T. Bartsch (2003)[3]. The last section will start by showing that spinors are isomorphic to quaternions, which explains the strong similarities that exist between both regularizations. We will then continue by studying these similarities in detail and comparing both regularizations in general. We will conclude by evaluating both regularizations and coming to a recommendation which one is the more practical regularization to use.

2 The Kepler problem

The first part of this paper will describe the Kepler problem. However, before defining the Kepler problem, we will begin by stating the well-known Newton's laws of motion. From hereon, we will use overhead arrows to denote vectors throughout the whole paper. The basic concepts of Multivariable Calculus we assume to be know to the reader are outlined in *Appendix 2: Multivariable Caluculus.*

2.1 Newton's Law's Of Motion

Newton's Laws of motion form the foundation for classical mechanics and therefore will play a very important role when considering the Kepler problem. Newton's first law of motion states that an object's velocity cannot change without a force acting upon it and can be stated as[1]:

Axiom 1. Newton's First Law: When viewed in an inertial reference frame, an object either is at rest or moves at a constant velocity, unless acted upon by a force.

Newton's second law of motion explains how the acceleration of an object relates to its mass and the forces acting upon it [1]:

Axiom 2. Newton's Second Law When viewed in an inertial reference frame, the acceleration $\vec{a}$ of a body is directly proportional to, and in the same direction as, the net force $\vec{F}$ acting on the body, and inversely proportional to its mass m, such that $\vec{F} = m\vec{a}$.

Newton's third law of motion explains how forces always appear together with a counterpart [1]:

Axiom 3. Newton's Third Law When a body A exerts a force on a second body B, the second body B simultaneously exerts a force equal in magnitude and opposite in direction to that of the first body A. So, if $\vec{F_A}$ is the force acting upon body A, as exerted by body B and $\vec{F_B}$ is the force acting on B as exerted by A, then $\vec{F_A} = -\vec{F_B}$

2.2 The original Kepler problem

The original Kepler problem described the orbital motion of the planets in our solar system around the sun. The Kepler problem can in general be used to analyze the motion of two (celestial) bodies, interacting through a gravitational force. For this situation the force one body exerts on the other body equals:

$$\vec{F}(\vec{r}) = m\ddot{\vec{r}} = \frac{-\mu m_1 m_2}{\vec{r}^2}\hat{\vec{r}}, \tag{1}$$

where μ is the gravitational constant, m_1 and m_2 are the masses of the two bodies and $\vec{r}$ is the seperation vector between the two bodies, pointing *towards* the body the force acts on (the hat over $\vec{r}$ denotes a unit vector).

Specifically for one of the bodies, body 1, equation (1) reads:

$$\vec{F}(\vec{r}) = m_1\ddot{\vec{r}} = \frac{-\mu m_1 m_2}{\vec{r}^2}\hat{\vec{r}}, \tag{2}$$

$$\ddot{\vec{r}} = \frac{-\mu m_2}{\vec{r}^2}\hat{\vec{r}}. \tag{3}$$

We can normalize equation (3) to give:

$$\ddot{\vec{r}} = -\frac{1}{\vec{r}^2}\hat{\vec{r}} = -\frac{\vec{r}}{|\vec{r}|^3}. \tag{4}$$

I will now illustrate how such an normalization can be achieved. For simplicity, let us assume that $\vec{r}$ is taken in units of *meters*, the standard unit of length (other units of length that are more common in celestial mechanics can be easily converted to meters). Now we choose our normalizing unit η as:

$$\eta = \frac{meters}{\sqrt{\mu m_2}}, \tag{5}$$

so 1 $meter$ = 1 $\sqrt{\mu m_2}\eta$ (because both the masses and the gravitational constant are positive, $\sqrt{\mu m_2}$ is real). If we now substitute the unit of $\vec{r}$ (*meters*) by $\sqrt{\mu m_2}\eta$ in equation (4) we obtain:

$$\ddot{\vec{r}} = -\frac{(\mu m_2)^{\frac{3}{2}}\vec{r}}{|\sqrt{\mu m_2}\vec{r}|^3} = -\frac{\vec{r}}{|\vec{r}|^3}, \tag{6}$$

where $\vec{r}$ is now in units of η.

Aside from its application in celestial mechanics, the Kepler problem has other applications as well, an important one being the electrostatic interaction between two charged particles. The Kepler problem can describe the orbital motion of two electrically charged particles interacting through an electrostatic force. Namely, this force obeys Coulomb's law, which states that the strength of the force is inversely related to the square of the distance between the two particles.

2.3 The Classical Kepler Problem

The general Kepler problem treats the motion of two bodies that interact through an attractive or repulsive force. However, first we will consider a simplified version of the general Kepler problem, which is sometimes considered the Classical Kepler problem: In the general Kepler problem both bodies can have *any* mass, but we will start by considering the case where one body has a *much* larger mass than the other body's mass. So, let us assume $m_A \gg m_B$, where m_A and m_B are the masses of bodies A and B, respectively. Let us now consider the consequences of this assumption according to Newton's Laws: According to Newton's Second Law (axiom 2), the acceleration $\vec{a}_A$ of a body A, as resulting from the force F_A is inversely proportional to its mass m_a:

$$\vec{a}_A = \frac{\vec{F}_A}{m_A}. \tag{7}$$

So the acceleration of body A decreases as m_A increases. Using Newton's Third Law (axiom 3) and Newton's Second Law (axiom 2) for body B gives us the relation between the acceleration and masses of the two bodies:

$$\vec{F}_A = -\vec{F}_B, \tag{8}$$

$$\vec{F}_b = m_B\vec{a}_B, \tag{9}$$

$$\vec{a}_A = -\frac{m_B}{m_A}\vec{F}_B, \tag{10}$$

$$\vec{a}_B = -\frac{m_A}{m_B}\vec{F}_A. \tag{11}$$

So $|\vec{a_A}| \ll |\vec{a_B}|$ for $m_A \gg m_B$. Therefore, assuming that $m_A \gg m_B$ allows us to ignore the acceleration of the much heavier body A, such that we may take body A as the origin of an inertial frame of reference. According to Newton's First Law (axiom 1), we may then describe the acceleration a_B of body B as a result of the force F_B. For now, we will forget about the heavy body A at the origin, the *central* body, and focus only on the force field it exerts. Thus, we can consider body B, to which we will refer as the *orbital* body, as in a gravitational force field. In this gravitational field the force F_B exerted on the orbital body is always oriented directly towards or away from the origin of the system, where the *central* body is located. We can read the equation of motion of the orbital body from equation (3):

$$\ddot{\vec{x}}_o = -\frac{\mu m_c \vec{x}_o}{|\vec{x}_o|^3}, \tag{12}$$

where $\vec{x}_o$ is the position vector of the orbital particle with respect to the inertial reference frame with the central body at the origin and m_c is the mass of the central body.

In section 2.4 we will consider important properties of force fields, which we will then apply to our orbital body.

2.4 Conservative Force Fields

We will now to consider force fields, such as the force field the orbital body is located in. With the force field, we will associate a potential, given by a smooth function $U : \Re^n \mapsto \Re$, for which

$$F(\vec{x}) = -\left(\frac{\partial U}{\partial x_1}(\vec{x}), ..., \frac{\partial U}{\partial x_n}(\vec{x})\right) \tag{13}$$

$$= -grad\ U(\vec{x}), \tag{14}$$

where $\vec{x}$ is the position vector. Any force field with which you can associate such a potential U is called *conservative*. The differential equations of motions of the orbital particle in the conservative force field are given by

$$\dot{\vec{x}} = \vec{v}, \tag{15}$$

$$\dot{\vec{v}} = \vec{a} = -\frac{1}{m}F(\vec{x}) = -\frac{1}{m}grad\ U(\vec{x}). \tag{16}$$

The kinetic energy K of the orbital body is given by

$$K = \frac{1}{2}m|\vec{v}|^2 \tag{17}$$

and the total energy is given by the sum of the kinetic energy and the potential

$$E = \frac{1}{2}m|\vec{v}|^2 + U(\vec{x}). \tag{18}$$

We can show that the energy of the orbital particle is constant for any solution curve of the equations of motion in the force field

$$\dot{E} = \frac{d}{dt}\left(\frac{1}{2}m|\vec{v}|^2 + U(\vec{x})\right). \tag{19}$$

Using the chain rule we have

$$\frac{d}{dt}|\vec{v}|^2 = \frac{d}{dt}\sqrt{v_1^2 + ... + v_n^2}^2 = \left(\frac{d}{dv_1}v_1^2\frac{dv_1}{dt} + ... + \frac{d}{dv_n}v_n^2\frac{dv_n}{dt}\right) = 2\vec{v}\cdot\dot{\vec{v}}, \tag{20}$$

$$\frac{d}{dt}U(\vec{x}) = (grad\ \ U)\cdot\dot{\vec{x}}, \tag{21}$$

$$\dot{E} = m\vec{v}\cdot\dot{\vec{v}} + (grad\ \ U)\cdot\dot{\vec{x}}. \tag{22}$$

Substituting equation (16) for $\dot{\vec{v}}$ gives:

$$\dot{E} = \vec{v}\cdot(-grad\ \ U) + (grad\ \ U)\cdot\vec{v} = 0, \tag{23}$$

so the energy is of the orbital body is constant through time.

2.5 Central Force Fields

As was discussed before, the orbital body is located in a force field, where the force $\vec{F}(\vec{x})$ always points towards or away from the origin for any position $\vec{x}$ (except for at the origin, which we will consider later). Such a force field is called *central.* Thus, we can write the force $\vec{F}(\vec{x})$ as a scalar multiple of the position vector $\vec{x}$, for any position $\vec{x}$:

$$\vec{F}(\vec{x}) = \lambda(\vec{x})\vec{x}, \tag{24}$$

where $\lambda(\vec{x})$ denotes a scalar coefficient that depends on $\vec{x}$. For our force field, $\lambda(\vec{x})$ is constant on a sphere $|\vec{x}| = constant$, because the field is conservative; it only depends on $r = |\vec{x}|$. Thus, for a conservative central force field, we have:

$$\vec{F}(\vec{x}) = \lambda(|\vec{x}|)\vec{x}. \tag{25}$$

This leads to the following equivalence relation:

Theorem 2.1. For a conservative vector field F, the following statements are equivalent:

1. *$\vec{F}$ is central*
2. $\vec{F} = f(|\vec{x}|)\vec{x}$
3. $\vec{F}(\vec{x}) = -grad\ \ U(\vec{X})$ *and* $U(\vec{x}) = g(|\vec{x}|)$ *for some function* g

Proof: see proof 6.1 in *Appendix 1: Proofs.*

Using the theorem 2.1 we can derive an interesting property of the motion of the orbital body in the central force field:

Theorem 2.2. The orbital body moving in the central force field always moves in a fixed plane containing the origin.

Proof: To show this, we consider the time derivative of the cross product of the orbital body's position and velocity vectors. Using equation (7.6) we obtain:

$$\frac{d}{dt}(\vec{x} \times \vec{v}) = \dot{\vec{x}} \times \vec{v} + \vec{x} \times \dot{\vec{v}} \tag{26}$$

$$= \vec{v} \times \vec{v} + \vec{x} \times \vec{a}. \tag{27}$$

$\vec{v} \times \vec{v}$ is obviously zero and, using theorem 2.1, we get $\vec{a} = F(\vec{x})/m = \frac{f(|\vec{x}|)}{m}\vec{x}$, so $\vec{a}$ is a scalar multiple of $\vec{x}$ and $\vec{x} \times \vec{a} = 0$. Thus

$$\frac{d}{dt}(\vec{x} \times \vec{v}) = 0. \tag{28}$$

■

It is equivalent to theorem 2.2 to state that the *angular momentum* is constant, as angular momentum is defined as $m(\vec{x} \times \vec{v})$. Therefore, the property of the motion of the orbital body in the central forcefield outlined in theorem 2.2 is often referred to as *conservation of angular momentum.*

2.6 The Kepler Central Force Field

We will now specify the force that is exerted on the orbital body, in order to make the force field specific to the Kepler problem. The force through which the orbital and central body interact is called a *central* force; its magnitude depends only on the distance between the two bodies and is directed along their separation vector. For the Kepler problem, this force obeys the *inverse square law*; the magnitude of the force is inversely proportional to the distance between the two bodies. In theory, this force could be both attractive or repulsive. However, in this paper we will only consider attractive forces, as in most "interesting" Kepler problems the force is attractive (with no other forces present, a repulsive kepler force will let two bodies only approach each other once until they fly off in different directions). With the central body at the origin, the force exerted on the orbital body by the origin is related to its position $\vec{x}$ by:

$$\vec{F} = -\frac{C}{|\vec{x}|^2}\hat{\vec{x}}, \tag{29}$$

where $C > 0$ is a constant. The exact relation is not specified for the Kepler problem; it depends on the nature of the force between the two bodies. For example, for two celestial bodies the force between them is a gravitational force and for two electrically charged particles the force is an electrostatic (Coulomb) force. We can choose units, depending on this nature of the force, such that constants are normalized. As we have seen, in the case of a gravitational force, the masses of the bodies and the gravitational parameter are normalized. From this we obtain:

$$\vec{F}(\vec{x}) = -\frac{1}{|\vec{x}|^2}\hat{\vec{x}} = -\frac{1}{|\vec{x}|^2}\frac{\vec{x}}{|\vec{x}|} = -\frac{\vec{x}}{|\vec{x}|^3} \tag{30}$$

and ecause of the normalization, we have that

$$\dot{\vec{v}} = -\frac{\vec{x}}{|\vec{x}|^3} \text{ and } \dot{\vec{x}} = \vec{v}. \tag{31}$$

As you can see, the differential equation (31) is not linear (see 6.2 in *Appendix 1: Proofs* if you do not readily see that) and has a singularity at the origin. This is because the force, and thus the acceleration become infinitely large at the origin. Physically, this makes sense, as the central body is located at the origin, so the orbital body cannot be at the origin as well. This is a non-physical situation, as it is impossible for both bodies to coexist at the origin independently and intact; they would collide. However, mathematically, it is preferable to remove this singularity at the origin, as it facilitates the numerical study of orbits near the origin. The methods of regularizing the Kepler problem, which we will discuss in the section to come, impose a regularization of equation (31) that will result in a linearized differential equation that has no singularity at the origin.

Consider now the potential

$$U(\vec{x}) = \frac{1}{|\vec{x}|}. \tag{32}$$

We can prove, using the potential (32), that the Kepler central force field is conservative. We do this by showing the force can be expressed in terms of the gradient of the potential, as in equation (24):

$$grad\ U(\vec{x}) = grad\left(-\frac{1}{|\vec{x}|}\right) = -\left(\frac{\partial|\vec{x}|^{-1}}{\partial x_1}, ..., \frac{\partial|\vec{x}|^{-1}}{\partial x_n}\right) \tag{33}$$

$$= \frac{1}{|\vec{x}|^2}\left(\frac{\partial|\vec{x}|}{\partial x_1}, ..., \frac{\partial|\vec{x}|}{\partial x_n}\right) \tag{34}$$

$$= \frac{1}{|\vec{x}|^2}\left(\frac{2x_1}{2\sqrt{x_1^2 + ... + x_n^2}}, ..., \frac{2x_n}{2\sqrt{x_1^2 + ... + x_n^2}}\right) \tag{35}$$

$$= \frac{\vec{x}}{|\vec{x}^3|} = -\vec{F}(\vec{x}). \tag{36}$$

Hence

$$\vec{F}(\vec{x}) = -grad\ U(\vec{x}). \tag{37}$$

■

Having found an expression for the potential, we can now specify the (total) energy as

$$E(\vec{x}, \vec{v}) = K(\vec{x}) + U(\vec{v}) = \frac{1}{2}|\vec{v}|^2 - \frac{1}{|\vec{x}|}. \tag{38}$$

Equation (38) is the energy integral of the Kepler force as specified in equation (30). *Proof:* See proof 6.4 in *Appendix 1: Proofs*.

2.7 The General Kepler Problem

So far, we have assumed one of the two bodies to be much heavier than the other, so that we could ignore the heavier body's acceleration and take it as the origin of an inertial frame of reference. This allowed us to reduce the two body problem to a one body problem; that of the orbital body in a conservative central force field (generated by the central body). However, for the general Kepler problem, there may be any mass ratio between the two bodies. We will now show that for the general Kepler problem, the two body problem can still be reduced to a one body problem, by considering the *center of momentum reference frame*. We will only demonstrate this for a gravitational Kepler force, as the procedure easily translates to other Kepler forces, such as the Coulomb force.

Consider two bodies who interact through a gravitational force with gravitational parameter μ. Let $\vec{x_1}$ and $\vec{x_2}$ be two position vectors, denoting the position of body 1 and body 2 respectively, with respect to some inertial frame of reference. Body 1 has mass m_1 and body 2 has mass m_2. From Newton's law of gravitation (equation (6)), the equations of motion of both bodies are given by:

$$m_1\ddot{\vec{x}}_1 = -\mu m_1 m_2 \frac{\vec{x_1} - \vec{x_2}}{|\vec{x_1} - \vec{x_2}|^3} \tag{39}$$

and

$$m_2\ddot{\vec{x}}_2 = -\mu m_1 m_2 \frac{\vec{x_2} - \vec{x_1}}{|\vec{x_2} - \vec{x_1}|^3}. \tag{40}$$

The *center of mass* coordinate is given by:

$$\vec{x}_{cm} := \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{m_1 + m_2}. \tag{41}$$

Thus,

$$\dot{\vec{x}}_{cm} = \frac{m_1\dot{\vec{x}}_1 + m_2\dot{\vec{x}}_2}{m_1 + m_2} \tag{42}$$

and

$$\ddot{\vec{x}}_{cm} = \frac{m_1\ddot{\vec{x}}_1 + m_2\ddot{\vec{x}}_2}{m_1 + m_2}. \tag{43}$$

A *center of momentum frame* is an inertial frame of reference in which the total momentum of the system is zero. By defining our system with respect to the center of mass frame we can restrict ourselves to the "interesting" motion of the two bodies relative to each other. The center of mass frame is simultaneously a center of momentum frame, since:

$$M_{tot}\dot{\vec{x}}_{cm} = (m_1 + m_2)\frac{m_1\dot{\vec{x}}_1 + m_2\dot{\vec{x}}_2}{m_1 + m_2} = m_1\dot{\vec{x}}_1 + m_2\dot{\vec{x}}_2 = p_{tot}, \tag{44}$$

where $M_{tot}\dot{\vec{x}}_{cm}$ is the momentum of the center of mass and p_{tot} is the total momentum of the system. We define the position of the bodies with respect to the center of mass frame as $\vec{x}_{cm1} = \vec{x}_1 - \vec{x}_{cm}$ and $\vec{x}_{cm2} = \vec{x}_2 - \vec{x}_{cm}$. The separation vector $\vec{r}$ between the two bodies is given by $\vec{r} = \vec{x_1} - \vec{x_2}$ and equations (39) and (40) in terms of the separation vector become:

$$m_1\ddot{\vec{x}}_1 = -\mu m_1 m_2 \frac{\vec{r}}{|\vec{r}|^3} \tag{45}$$

and

$$m_2\ddot{\vec{x}}_2 = \mu m_1 m_2 \frac{\vec{r}}{|\vec{r}|^3}. \tag{46}$$

Now we multiply equation (46) by m_1 and substract it from equation (45), multiplied by m_2 to get

$$m_1 m_2(\ddot{\vec{x}}_1 - \ddot{\vec{x}}_2) = -\frac{\mu m_1 m_2(m_1+m_2)\vec{r}}{\vec{r}^3}, \tag{47}$$

$$\ddot{\vec{r}} = -\frac{\mu(m_1+m_2)\vec{r}}{\vec{r}^3}. \tag{48}$$

Thus, the two body general Kepler problem reduces to the one body classical Kepler problem, with the mass of the central body replaced by the total mass of the system and the position vector of the orbital body replaced by the separation vector. Note that for $m_1 \gg m_2$ equation (48) reduces to equation (12), with body 2 as the orbital body. Namely, with body 1, the central body, at the origin of the reference frame, the separation vector $\vec{r}$ between the two bodies is equal to the position vector $\vec{x}_o$ of the orbital body and for $m_1 \gg m_2$ we have $m_1 + m_2 \approx m_1$.

Now, all we need to know is how to solve the classical Kepler problem to obtain $\vec{r}$, from which we can then find the trajectories of the individual bodies as follows: From equation (41) and $\vec{r} = \vec{x}_1 - \vec{x}_2$ we have:

$$m_1\vec{x}_1 + m_2(\vec{x}_1 - \vec{r}) = (m_1+m_2)\vec{x}_c m, \tag{49}$$

$$\vec{x}_1 = \vec{x}_c m + \frac{m_2\vec{r}}{m_1+m_2}, \tag{50}$$

and analogously:

$$\vec{x}_2 = \vec{x}_c m - \frac{m_1\vec{r}}{m_1+m_2}, \tag{51}$$

where the first part of the right hand side represents the location of the center of mass and the second part the location of the body with respect to the center of mass.

Thus, we have shown that the general Kepler problem for two bodies interacting through a gravitational force with gravitational constant μ can be solved by reducing the problem to a system of classical Kepler problems. Therefore, we will focus our attention mainly on the classical Kepler problem for the remainder of this paper, as the general Kepler problem follows naturally from it.

2.8 Regularizing the two dimensional Kepler problem

We will now discuss the method of regularizing the classical Kepler problem in two dimensions as introduced by Levi-Civita (1920) [7]. Even though this method only considers two dimensions, it is still very relevant for physical cases, as the conservation of angular momentum restricts a three dimensional (unpertubed) Kepler problem to a (two dimensional) plane and thus, the method introduced by Levi-Civita can be used to find the orbit of a body in a central force field in three-dimensional space. The method uses complex numbers to express two-dimensional vectors. Thus, a vector $\vec{x} = (x_1\vec{\sigma}_1, x_2\vec{\sigma}_2)$, where $\vec{\sigma}_1$ and $\vec{\sigma}_2$ are two orthogonal unit vectors that make up two-dimensional space, is represented by the complex number $\tilde{\mathbf{x}} = x_1 + ix_2 \in \mathbb{C}$ (complex numbers will be denoted in bold-face lower case letters) . We will not prove any of the relations given here, as the aim of this section is only to illustrate Levi-Civita's method of regularizing the Kepler problem.

First, the physical time t is replaced by the fictitious time τ, according to the following relation:

$$dt = r d\tau, \quad \frac{d}{d\tau} = (.)'. \tag{52}$$

Thus, the rate of change $\frac{dt}{dt} = r$ is made proportional to the distance r to the singularity at the origin. Therefore, the closer to the origin, the less dt increases when $d\tau$ increases; the physical time runs in *slow motion* close to the origin with respect to the fictitious time τ. Using product- and chain rules, we obtain the following relations between differentiation with respect to t and τ:

$$\frac{d}{dt} = \frac{1}{r}\frac{d}{d\tau}, \quad \frac{d^2}{dt^2} = \frac{1}{r^2}\frac{d^2}{d\tau^2} - \frac{r'}{r^3}\frac{d}{d\tau}. \tag{53}$$

Using these relations, we can write equation (6) as

$$r\mathbf{x}'' - r'\mathbf{x}' + \mu\mathbf{x} = r^3. \tag{54}$$

Now, the complex representation $\mathbf{x}$ of the physical coordinate $\vec{x}$ is represented as the square $\mathbf{u^2}$ of a complex variable $\mathbf{u}= u_1 + iu_2 \in \mathbb{C}$

$$\mathbf{x} = \mathbf{u^2}. \tag{55}$$

From equation (55) we obtain

$$r = |\mathbf{x}| = |\mathbf{u}|^{\mathbf{2}} = \mathbf{u}\bar{\mathbf{u}}, \tag{56}$$

where $\bar{\mathbf{u}}$ is the complex conjugate of $\mathbf{u}$. Differentiation with respect to τ of equations (55) and (56) yields:

$$\mathbf{x}' = \mathbf{2uu}', \quad \mathbf{x}" = \mathbf{2}\left(\mathbf{uu}" + \mathbf{u'^2}\right), \quad r' = \mathbf{u}'\bar{\mathbf{u}} + \mathbf{u}\bar{\mathbf{u}}'. \tag{57}$$

Substituting equation (57) into equation (54) yields (after algebraic manipulation):

$$2r\mathbf{u}^{''} + (\mu - 2|\mathbf{u}^{'}|^2)\mathbf{u} = r^2\bar{\mathbf{u}}. \tag{58}$$

Now, using equations (38), (52) and (57) we obtain:

$$\dot{\mathbf{x}} = \frac{1}{r}\mathbf{2uu}', \quad \frac{1}{2}|\dot{\mathbf{x}}|^2 = 2\frac{|\mathbf{u}'|^2}{r} \tag{59}$$

and we express equation (38) in terms of the derivative with respect to τ:

$$\mu - 2|\mathbf{u}'|^2 = rh. \tag{60}$$

Finally, substituting equations (59) and (60) into equation (58) yields

$$2\mathbf{u}^{''} + h\mathbf{u} = 0. \tag{61}$$

Thus, we have obtained a linear differential equation without the singularity at the origin. The general solution is given by

$$\mathbf{u} = A\cos\omega\tau + iB\sin\omega\tau, \qquad \omega = \sqrt{h/2}, \tag{62}$$

which parametrizes an ellipse with the origin at the center and which we will square ($\mathbf{u}^2$) to obtain an ellipse which has the origin as one of its foci (see figure 1):

$$\mathbf{x} = \mathbf{u^2} = \frac{A^2 - B^2}{2} + \frac{A^2 + B^2}{2}\cos 2\omega\tau + iAB\sin 2\omega\tau. \tag{63}$$

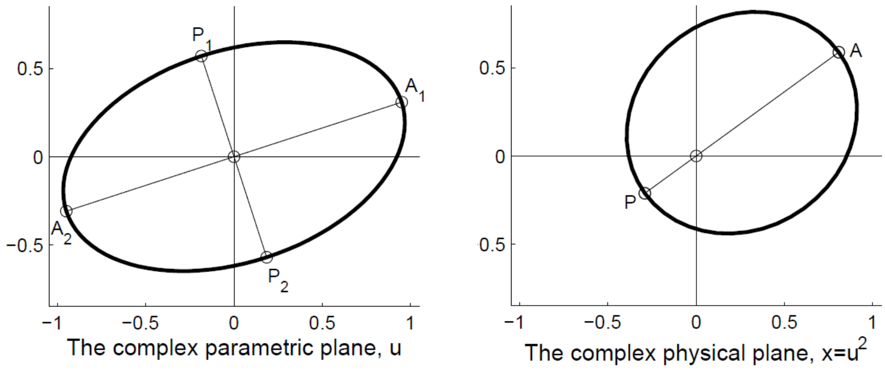

Figure 1: Origin Centered and (Squared) Origin Focused Ellipse

In *Appendix 3: The resulting orbital motion from the Kepler problem* we will explain how this describes an elliptic motion of the orbital body.

2.9 The Kustaanheimo-Stiefel transformation

Before moving on to the main part of this paper, we need to introduce one more concept: In three dimensions, introducing the square root coordinate $\mathbf{u}$ as used in the Levi-Civita method (equation 55), is replaced by the Kustaanheimo-Stiefel transformation. The Kustaanheimo-Stiefel transformation, or KS transformation, was first introduced by P. Kustaanheimo and E. Stiefel (1965) [6]. The KS transformation is a transformation that maps four dimensions into three dimensions. The general form of this transformation is given by: $\Re^4 \mapsto \Re^3 : (u_0, u_1, u_2, u_3) \mapsto (x_0, x_1, x_2)$ defined as

$$\begin{aligned} x_0 &= u_0^2 - pu_1^2 - pqu_3^2 \\ x_1 &= 2(u_0u_2 + pu_1u_3) \\ x_2 &= 2(u_0u_3 + u_1u_2), \end{aligned}$$

where p and q are equal to ± 1. P. Kustaanheimo and E. Stiefel proposed a version of the KS transformation with $p = q = -1$, which gives:

$$\begin{aligned} x_0 &= u_0^2 + u_1^2 - u_2^2 - u_3^2 \\ x_1 &= 2(u_0u_2 - u_1u_3) \\ x_2 &= 2(u_0u_3 + u_1u_2). \end{aligned}$$

3 Regularizing the Kepler problem in three dimensions

Even though the conservation of angular momentum reduces the Kepler problem in three dimensions to a two-dimensional problem, cases do exist where additional forces are present. Thus, the angular momentum of the orbital body around the central body is not conserved and the orbit therefore cannot be contained in a two dimensional plane. The orbital body's trajectory is then extended to a three-dimensional motion. For example, perturbing the Kepler problem for the orbit of an electron around a nucleus by an electric or magnetic field constitutes such a case. Studying such cases is beyond the scope of this paper, because we want to compare the quaternion method to the spinor method at the most basic level for solving the Kepler problem, namely regularizing the *unperturbed* Kepler problem. However, the existence of such cases does constitute the reason for needing to have a formalism that can express the Kepler problem in three dimensions.

3.1 Regularizing the Kepler problem using quaternions

This section will show how the Kepler problem can be solved using quaternion algebra, according to the paper by J. Walvogel (2007)[13]. We will begin by giving a summary of quaternion algebra and then move to solving the Kepler problem.

3.1.1 Quaternion algebra

Quaternions were first described by Hamilton. Complex numbers can be interpreted as points in a two-dimensional plane and Hamilton was trying to extend the concept of complex numbers by finding a "complex" representation of points in three-dimensional space. As coordinates in three-dimensional space are triples of numbers, Hamilton needed to establish how to calculate with triples. However, he could not find a way to do so, until he realized he needed a fourth dimension for the purpose calculating with triples [2].

Quaternions extend the complex plane to four dimensions by introducing the additional "imaginary numbers" j and k. Quaternion space allows for the expression of vectors in four dimensions, in units of the real unit 1 and the imaginary units i, j, and k. The imaginary units i, j and k obey the following relation

$$i^2 = j^2 = k^2 = ijk = -1. \tag{64}$$

These imaginary units do not commute, such that $ij \neq ji$. By requiring commutation with the real number -1, we can obtain the following multiplication rules from equation (64):

$$(ijk)k = (-1)k, \tag{65}$$

$$ij(-1) = (-1)k, \tag{66}$$

$$ij = k \tag{67}$$

and

$$i(ij) = i(k), \tag{68}$$

$$ik = -j. \tag{69}$$

Thus, $ij = k$ and $ik = -j$. The quaternion space's imaginary units also obey a cyclic permutation $i \to j \to k \to i$

$$ijk = kij = -1, \tag{70}$$

$$kij = k(ijk)/k = k(-1)/k = (-1)k/k = -1 = ijk = jki. \tag{71}$$

We can summarize these relations as

$$ij = -ji = k, \tag{72}$$

$$jk = -kj = i, \tag{73}$$

$$ki = -ik = j. \tag{74}$$

Now we can define quaternions: A quaternion is an object expressed in the imaginary units i, j and k and the real unit 1. Given the real numbers $u_l \in \Re$, $l = 0, 1, 2, 3$ a quaternion, denoted with a lower-case bold-face letter, $\mathbf{u} \in \mathbb{U}$, where $\mathbb{U}$ denotes the set of all quaternions, can be expressed as:

$$\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3. \tag{75}$$

The quaternion part consists of a real part, u_0, and a *quaternion part* $iu_1 + ju_2 + ku_3$. Quaternion algebra consists of the multiplications as stated in equations (72), (73) and (74) and vector space addition. In general multiplication in non-commutative, as discussed above. However, any quaternion commutes with real numbers:

$$c\mathbf{u} = \mathbf{u}c, \tag{76}$$

with $c \in \Re$ and $\mathbf{u} \in \mathbb{U}$. Also, the associative law holds:

$$(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw}). \tag{77}$$

Proof: see proof 6.6 in *Appendix 1: Proofs*.

The quaternions form a ring: The quaternions, together with quaternion addition (which is associative and commutative) form an abelian group. This abelian group, together with the second group operation, quaternion multiplication (which is associative and distributive), form a ring. However, because quaternion multiplication is not commutative, the quaternions do not form a field as the complex numbers do.

As stated before, quaternions can be used to represent four-dimensional vectors. A quaternion $\mathbf{u}$ may be associated with the corresponding four-dimensional vector $u = (u_0, u_1, u_2, u_3) \in \Re^4$. We can also associate two different kinds of particular quaternions with a three-dimensional real vector. A *pure quaternion* is a quaternion whose real part equals zero and can thus be expressed *purely* in the quaternion imaginary units i, j and k. Such a pure quaternion $\mathbf{u} = iu_1 + ju_2 + ku_3$ can be used to represent the real three-dimensional vector $\vec{u} = (u_1, u_2, u_3) \in \Re^3$. The other type of quaternion that can be associated with a real three-dimensional vector is a quaternion whose k-component is equal to zero (which is equivalent to some quaternion with vanishing i or j part because of the cyclic property of imaginary quaternion units). Such a quaternion $\mathbf{u} = u_0 + iu_1 + ju_2$ can be expressed as a three-dimensional real vector $\vec{u} = (u_0, u_1, u_2) \in \Re^3$. We define the *conjugate* $\bar{\mathbf{u}}$ of the quaternion $\mathbf{u}$ by changing the sign of the imaginary parts (the real part and the magnitude of the imaginary parts remain unchanged) as

$$\bar{\mathbf{u}} := u_0 - iu_1 - ju_2 - ku_3. \tag{78}$$

The *modulus* $|\mathbf{u}|$ of $\mathbf{u}$ we define as

$$|\mathbf{u}|^2 = \mathbf{u}\bar{\mathbf{u}} = \bar{\mathbf{u}}\mathbf{u} = (u_0 + iu_1 + ju_2 + ku_3)(u_0 - iu_1 - ju_2 - ku_3) \tag{79}$$

$$= u_0^2 - iu_0u_1 - ju_0u_2 - ku_0u_3 + iu_1u_0 + u_1^2 - iju_1u_2 - iku_1u_3 + ... \tag{80}$$

$$... + ju_2u_0 - jiu_2u_1 + u_2^2 - jku_2u_3 + ku_3u_0 - kiu_3u_1 - kju_3u_2 + u_3^2$$

$$= \Sigma_{l=0}^3 u_l^2 - iju_1u_2 - iku_1u_3 - jiu_2u_1 - jku_2u_3 - kiu_3u_1 - kju_3u_2, \tag{81}$$

$$|\mathbf{u}|^2 = \Sigma_{l=0}^3 u_l^2. \tag{82}$$

Thus, the *modulus* $|\mathbf{u}|$ of $\mathbf{u}$ is given by $|\mathbf{u}| = \sqrt{|\mathbf{u}|^{\mathbf{2}}} = \sqrt{\Sigma_{l=0}^3 u_l^2}$. Conjugating a product of quaternions reverses their order, as with transposing of a matrix product:

$$\overline{\mathbf{u}\mathbf{v}} = \bar{\mathbf{v}}\bar{\mathbf{u}}. \tag{83}$$

Proof: see proof 6.7 in *Appendix 1: Proofs.*

Dividing by quaternions is defined as either left- or right-multiplying by the inverse, as common with matrix algebra. The inverse is defined as

$$\mathbf{u}^{-\mathbf{1}} = \frac{\bar{\mathbf{u}}}{\mathbf{u}\bar{\mathbf{u}}}. \tag{84}$$

Thus, division by a quaternion $\mathbf{u} \neq 0$ is performed by left or right multiplication by $\mathbf{u}^{-\mathbf{1}} = \frac{\bar{\mathbf{u}}}{\mathbf{u}\bar{\mathbf{u}}} = \frac{\bar{\mathbf{u}}}{|\mathbf{u}|}$. Because it is possible to divide quaternions, they form a division ring (i.e. a ring in which division is possible) [2].

Quaternions can be used to represent rotations in $\Re^3$. Take $\vec{a} = (a_1, a_2, a_3) \in \Re^3$, where $\vec{a}$ is a unit vector, so $|\vec{a}| = 1$. We will use quaternions to rotate around this vector $\vec{a}$ by an angle ω. To do so, we define the unit quaternion

$$\mathbf{r} := \cos\frac{\omega}{2} + (ia_1 + ja_2 + ka_3)\sin\frac{\omega}{2}. \tag{85}$$

Thus,

$$|\mathbf{r}| = \sqrt{\cos^2\frac{\omega}{2} + (ia_1^2 + ja_2^2 + ka_3^2)\sin 2\frac{\omega}{2}} = \sqrt{\cos^2\frac{\omega}{2} + |\vec{a}|^2\sin^2\frac{\omega}{2}} = \sqrt{\cos^2\frac{\omega}{2} + \sin^2\frac{\omega}{2}} = 1, \tag{86}$$

so

$$\mathbf{r^{-1}} = \frac{\bar{\mathbf{r}}}{|\mathbf{r}|} = \bar{\mathbf{r}}. \tag{87}$$

Now, take $\vec{x} \in \Re^3$ to be an arbitrary vector and relate $\mathbf{x} = ix_1 + jx_2 + kx_3$ to the pure quaternion. We can now describe a right-handed rotation of $\vec{x}$ around the rotation vector $\vec{a}$ as its axis by an angle ω by the mapping:

$$\mathbf{x} \mapsto \mathbf{y} = \mathbf{rxr^{-1}} \tag{88}$$

The proof of this is beyond the scope of this paper, but can be found in the paper by J. Waldvogel (2006) [11]. We will also define the *star conjugate* $\mathbf{u}^*$ of the quaternion $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$:

$$\mathbf{u}^* := u_0 + iu_1 + ju_2 - ku_3. \tag{89}$$

We can obtain the *star conjugate* $\mathbf{u}^*$ of the quaternion $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$ from the regular conjugate $\bar{\mathbf{u}}$:

$$k\bar{\mathbf{u}}k^{-1} = -k\bar{\mathbf{u}}k = -k(u_0 - iu_1 - ju_2 - ku_3)k = -(u_0k - kiu_1 - kju_2 - kku_3)k \tag{90}$$

$$= -(u_0k - ju_1 + iu_2 + u_3)k = -(u_0kk - jku_1 + iku_2 + ku_3) = u_0 + iu_1 + ju_2 - ku_3 = \mathbf{u}^*. \tag{91}$$

We prove the following properties for future reference:

$$(\mathbf{uv})^* = \mathbf{v}^*\mathbf{u}^*. \tag{92}$$

Proof: See proof 6.8 in *Appendix 1: Proofs*. Furthermore,

$$(\mathbf{u}^*)^* = (u_0 + iu_1 + ju_2 - ku_3)^* = u_0 + iu_1 + ju_2 + ku_3 = \mathbf{u}, \tag{93}$$

and

$$|\mathbf{u}^*|^2 = (u_0)^2 + (u_1)2 + (u_2)^2 + (-u_3)^2 = (u_0)^2 + (u_1)2 + (u_2)^2 + (u_3)^2 = |\mathbf{u}|^2. \tag{94}$$

Consider the following mapping:

$$\mathbf{u} \in U \mapsto \mathbf{x} = \mathbf{uu}^* \tag{95}$$

According to the properties of the star conjugate we deduced, we see that

$$\mathbf{x}^* = (\mathbf{u}^*)^*\mathbf{u}^* = \mathbf{x}, \tag{96}$$

so $x_3 = -x_3 = 0$. Thus, $\mathbf{x}$ is a quaternion of the form $\mathbf{x} = x_0 + ix_1 + jx_2$, which is related to the vector $\vec{x} = (x_0, x_1, x_2) \in \Re^3$. Take $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$. By applying the mapping to $\mathbf{u}$, we obtain

$$\mathbf{x} = (u_0+iu_1+ju_2+ku_3)(u_0+iu_1+ju_2+ku_3)^* = (u_0+iu_1+ju_2+ku_3)(u_0+iu_1+ju_2-ku_3) \tag{97}$$

$$= u_0^2 + iu_0u_1 + ju_0u_2 - ku_0u_3 + iu_1u_0 - u_1^2 + ku_1u_2 + ju_1u_3... \tag{98}$$

$$...ju_2u_0 - ku_2u_1 - u_2^2 - iu_2u_3 + ku_3u_0 + ju_3u_1 - iu_3u_2 + u_3^2 \tag{99}$$

$$= u_0^2 - u_1^2 - u_2^2 + u_3^2 + i(2u_0u_1 - 2u_2u_3) + j(2u_0u_2 + 2u_1u_3). \tag{100}$$

Thus, for the vector $\underline{\mathrm{x}}= (x_0, x_1, x_2) \in \Re^3$ associated with this mapping, we obtain

$$\begin{aligned} x_0 &= u_0^2 - u_1^2 - u_2^2 + u_3^2 \\ x_1 &= 2u_0u_1 - 2u_2u_3 \\ x_2 &= 2u_0u_2 + 2u_1u_3. \end{aligned} \tag{101}$$

This resembles exactly the Kustaanheimo-Stiefel transformation.

Theorem 3.1. The Kustaanheimo-Stiefel transformation $u = (u_0, u_1, u_2, u_3) \in \Re^3 \mapsto \underline{\mathrm{x}}= (x_0, x_1, x_2) \in \Re^3$ is given by the quaternion relation

$$\mathbf{x} = \mathbf{u}\mathbf{u}^*,$$

where $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$, $\mathbf{x} = x_0 + ix_1 + jx_2$, and $\mathbf{u}^*$ is defined as in equation (89).

Additionally, we note for future use that

Lemma 3.2. $r := |\underline{\mathrm{x}}| = |\mathbf{u}|^\mathbf{2} = \mathbf{u}\bar{\mathbf{u}}$.

Proof:

$$|\underline{\mathrm{x}}|^2 = \mathbf{x}\bar{\mathbf{x}} = (\mathbf{u}\mathbf{u}^*)(\overline{\mathbf{u}\mathbf{u}^*}) = \mathbf{u}(\mathbf{u}^*\bar{\mathbf{u}}^*)\bar{\mathbf{u}} = |\mathbf{u}^*|^\mathbf{2}|\mathbf{u}|^\mathbf{2} = |\mathbf{u}|^\mathbf{4}. \tag{102}$$

We will now look at differentiation of the mapping as defined in equation (95). As the mapping (95) or (101) maps from $\Re^4$ to $\Re^3$, it leaves one degree of freedom in the parametric space undetermined. By imposing the "Bilinear relation":

$$2(u_3du_0 - u_2du_1 + u1du_2 - u_0du_3) = 0, \tag{103}$$

between the vector $\vec{u} = (u_0, u_1, u_2, u_3)$ and its differential du on orbits. The tangential mapping of equation (95) becomes a linear mapping with an orthogonal matrix. We will now show the consequence this has on the differentiation of equation (95). Since the quaternion product is non-commutative, the differential mapping of (95) is given by

$$d\mathbf{x} = d\mathbf{u} \cdot \mathbf{u}^* + \mathbf{u} \cdot d\mathbf{u}^*. \tag{104}$$

Furthermore, we can use equation 103 in the form of the commutation relation

$$\mathbf{u} \cdot d\mathbf{u}^* - d\mathbf{u} \cdot \mathbf{u}^* = 0, \tag{105}$$

to obtain

$$d\mathbf{x} = 2\mathbf{u} \cdot d\mathbf{u}^*. \tag{106}$$

Thus, imposing the bilinear equation (103) is equivalent with requiring that the tangential mapping of (95) behaves as in commutative algebra.

We will now try to find an expression for the inverse mapping of the mapping (95). However, because the mapping (95) does not preserve its four dimensions, but maps to three dimensions and leaves one degree of freedom, we cannot find a single well-defined inverse how we usually would. Therefore we will look at the *fiber* of the original $\Re^4$ space corresponding to the mapping (95) to find the set of quaternions $\mathbf{u}$ that are mapped by (95) to a certain $\mathbf{x} = \mathbf{u}\mathbf{u}^*$. Thus, given a certain quaternion $\mathbf{x} = x_0 + ix_1 + jx_2$, we want to find the *fiber* of this quaternion that corresponds to all quaternions $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$ for which $\mathbf{u}\mathbf{u}^*$ equals $\mathbf{x}$. To do so, we first find a particular solution $\mathbf{u} := \mathbf{v} = \mathbf{v}^* = v_o + iv_1 + jv_2$ which has a vanishing k-component as well. Because $\mathbf{v} = \mathbf{v}^*$ (because of the vanishing k-component) $\mathbf{v}\mathbf{v}^* = \mathbf{v^2}$, so we can derive $\mathbf{v}$ as one of the quaternion square roots of $\mathbf{x}$. Thus,

$$\mathbf{v} = \frac{\mathbf{x} + |\mathbf{x}|}{\sqrt{2(x_0 + |\mathbf{x}|)}}. \tag{107}$$

This is a formula for the square root of the complex number $\mathbf{x} = x_0 + ix_1 \in \mathbb{C}$ (with $\mathbb{C}$ the set of complex numbers). The *fiber* corresponding to $\mathbf{x}$, the entire set of quaternions $\mathbf{u}$ for which $\mathbf{u}\mathbf{u}^* = \mathbf{x}$, the *fiber* corresponding to $\mathbf{x}$ (which can geometrically be represented by a circle in $\Re^4$, parametrized by ϕ), is given by

$$\mathbf{u} = \mathbf{v} \cdot e^{k\phi} = \mathbf{v}(\cos\phi + k\sin\phi). \tag{108}$$

You check that equation (108) indeed satisfies $\mathbf{u}\mathbf{u}^* = \mathbf{x}$:

$$\mathbf{u}\mathbf{u}^* = (\mathbf{v}e^{k\phi})(\mathbf{v}e^{k\phi})^* = \mathbf{v}(e^{k\phi}e^{-k\phi})\mathbf{v}^* = \mathbf{v}\mathbf{v}^* = \mathbf{x}. \tag{109}$$

However, proving that all elements of the fiber corresponding to $\mathbf{x}$ are of this form is beyond the scope of this paper.

3.1.2 Solving the Kepler problem using Quaternions

In this section we will use the previously derived quaternion algebra to solve the Kepler problem. We will solve the Kepler problem by finding the equation of motion of the orbital body's trajectory around the central body by using four parameters: $(u_0, u_1, u_2, u_3) := u \in \Re^4$, with the corresponding quaternion $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$. As explained before, for a situation where the orbital body possesses an angular momentum with respect to the central body at the origin, and no additional forces are present, the orbital body moves in a plane spanned by the velocity vector of the orbital body and the separation vector between the orbital body and the central body at the origin. Since the motion is planar, we can consider the plane spanned by u_0 and u_1 and take u_2 and u_3 to be zero, such that $\mathbf{u} = u_0 + iu_1 \in U$. Solving this problem can be done by the Levi-Civita (1920) [7] method.

We begin with the equations of motion (31) and (38). Equations (38) and (30) govern the Kepler problem as discussed in the section 2.6. We have kept the gravitational factor μ instead of normalizing it out, in correspondence with the paper by Waldvogel (2007) [13]. In quaternion notation, equations (31) and (38) are given by:

$$\ddot{\mathbf{x}} + \mu\frac{\mathbf{x}}{r^3} = 0 \in U. \tag{110}$$

As before, t denotes the time, μ is the gravitational parameter, $r = |\vec{x}|$ and $(\dot{})$ denotes the derivative with respect to time of the parameter in the brackets. The position of the orbital particle is given by $\vec{x} = (x_0, x_1, x_2) \in \Re^3$ which corresponds to the quaternion $\mathbf{x} = x + 0 + ix_1 + jx_2 \in U$. The energy integral of (110) is given by:

$$\frac{1}{2}|\dot{\mathbf{x}}|^2 - \frac{\mu}{r} = -h = constant. \tag{111}$$

The minus sign in front of the h is chosen such that $h > 0$ corresponds to an elliptic orbit. We will now find the equations of motion of the trajectory of the orbital particle around the central body located at the origin.

First, we will introduce a new "time variable" τ. This independent time variable defines a *fictitious time* according to the Sundman (1970) [10] transformation:

$$dt := r \cdot d\tau. \tag{112}$$

From now on, we will denote differentiation with respect to the fictitious time variable τ as: $\frac{d}{d\tau}(\ ) = (\ )'$. Thus, the ratio $\frac{dt}{d\tau}$ of the two infinitesimal increments is, from the definition 112, proportional to the distance r to the origin of the system. Thus, the orbit runs in slow-motion as r becomes small. The vector equations (110) and (111) in terms of τ are transformed to:

$$r\mathbf{x}'' - r'\mathbf{x}' + \mu\mathbf{x} = 0, \tag{113}$$

$$\frac{1}{2r^2}|\mathbf{x}'|^2 - \frac{\mu}{r} = -h = constant. \tag{114}$$

The transformation of equations (114) and (113) was achieved as follows

$$\ddot{\mathbf{x}} = \frac{d}{dt}\left(\frac{d}{dt}x\right) = \frac{d}{dt}\left(\frac{dx}{d\tau}\frac{d\tau}{dt}\right) = \frac{d}{dt}\left(x'r^{-1}\right) \tag{115}$$

$$= r^{-1}\frac{dx'}{dt} + x'\frac{dr^{-1}}{dt} = r^{-1}\frac{dx'}{d\tau}\frac{d\tau}{dt} - r^{-2}x'\frac{dr}{d\tau}\frac{d\tau}{dt} = r^{-2}x'' - r^{-3}r'x'. \tag{116}$$

Substituting equation (116) for $\ddot{\mathbf{x}}$ in equation (113) yields

$$r^{-2}x'' - r^{-3}r'x' + \mu\frac{\mathbf{x}}{r^3} = 0 = rx'' - r'x' + \mu\mathbf{x} \tag{117}$$

and substituting $\dot{\mathbf{x}} = \mathbf{x}'r^{-1}$ yields

$$\frac{1}{2r^2}|\mathbf{x}'r^{-1}|^2 - \frac{\mu}{r} = -h = \frac{1}{2r^2}|\mathbf{x}'|^2 - \frac{\mu}{r}. \tag{118}$$

Now, using the previously derived differentiation rules, we will substitute the image of mapping (95) into the in fictitious time τ parametrized Kepler equations:

$$\mathbf{x} = \mathbf{u}\mathbf{u}^*, \tag{119}$$

$$r := |\mathbf{x}| = |\mathbf{u}|^{\mathbf{2}} = \mathbf{u}\bar{\mathbf{u}}. \tag{120}$$

Using equation (106), we can derive the following relations

$$\mathbf{x}' = \frac{d\mathbf{x}}{d\tau} = 2\mathbf{u} \cdot \frac{d\mathbf{u}^*}{d\tau} = 2\mathbf{u}\mathbf{u}^{*'}, \tag{121}$$

$$\mathbf{x}'' = \frac{d}{d\tau}\mathbf{x}' = \frac{d}{d\tau}(2\mathbf{u}\mathbf{u}^{*'}) = 2\frac{d}{d\tau}(\mathbf{u}^{*'}) \cdot \mathbf{u} + 2\frac{d}{d\tau}(\mathbf{u}) \cdot \mathbf{u}^{*'} = 2\mathbf{u}\mathbf{u}^{*''} + 2\mathbf{u}'\mathbf{u}^{*'} \tag{122}$$

and

$$r' = \frac{dr}{d\tau} = \frac{d}{d\tau}(\mathbf{u}\bar{\mathbf{u}}) = \frac{d}{d\tau}(\mathbf{u})\bar{\mathbf{u}} + \mathbf{u}\frac{d}{d\tau}(\bar{\mathbf{u}}) = \mathbf{u}'\bar{\mathbf{u}} + \mathbf{u}\bar{\mathbf{u}}'. \tag{123}$$

Substituting equations (121), (122) and (123) into the fictitious time parametrized first Kepler equation (113) yields:

$$(\mathbf{u}\bar{\mathbf{u}})(2\mathbf{u}\mathbf{u}^{*''} + 2\mathbf{u}'\mathbf{u}^{*'}) - (\mathbf{u}'\bar{\mathbf{u}} + \mathbf{u}\bar{\mathbf{u}}')2\mathbf{u}\mathbf{u}^{*'} + \mu\mathbf{u}\mathbf{u}^* = 0. \tag{124}$$

We can use the distributive law to eliminate the second and third term:

$$(\mathbf{u}\bar{\mathbf{u}})2\mathbf{u}'\mathbf{u}^{*'} - (\mathbf{u}'\bar{\mathbf{u}})2\mathbf{u}\mathbf{u}^{*'} = 2(\mathbf{u}\bar{\mathbf{u}})\mathbf{u}'\mathbf{u}^{*'} - 2\mathbf{u}'(\bar{\mathbf{u}}\mathbf{u})\mathbf{u}^{*'} = 2r(\mathbf{u}'\mathbf{u}^{*'} - \mathbf{u}'\mathbf{u}^{*'}) = 0. \tag{125}$$

Furthermore, equation (105) can be used to derive:

$$\frac{d}{d\tau}(\mathbf{u} \cdot d\mathbf{u}^*) = \frac{d}{d\tau}(d\mathbf{u} \cdot \mathbf{u}^*), \tag{126}$$

$$\mathbf{u} \cdot \frac{d\mathbf{u}^*}{d\tau} = \mathbf{u}\mathbf{u}^{*'} = \frac{d\mathbf{u}}{d\tau} \cdot \mathbf{u}^* = \mathbf{u}'\mathbf{u}^*. \tag{127}$$

Now we use the relation $\mathbf{u}\mathbf{u}^{*'} = \mathbf{u}'\mathbf{u}^*$ to simplify the fourth term of equation (124)

$$-(\mathbf{u}\bar{\mathbf{u}}')2\mathbf{u}\mathbf{u}^{*'} = -2\mathbf{u}(\bar{\mathbf{u}}'\mathbf{u}')\mathbf{u}^* = -2|\mathbf{u}'|^{\mathbf{2}}\mathbf{u}\mathbf{u}^*. \tag{128}$$

This turns equation (124) into

$$2(\mathbf{u}\bar{\mathbf{u}})(\mathbf{u}\mathbf{u}^{*''}) - 2|\mathbf{u}'|^{\mathbf{2}}\mathbf{u}\mathbf{u}^* + \mu\mathbf{u}\mathbf{u}^* = 0 \tag{129}$$

and by making use of $\mathbf{u}\bar{\mathbf{u}} = r$ and left-dividing by $\mathbf{u}$, we obtain

$$2r\mathbf{u}^{*''} + (\mu - 2|\mathbf{u}'|^2)\mathbf{u}^* = 0. \tag{130}$$

Using respectively equations (79), (121) and (317) we derive the following

$$|\mathbf{x}'|^2 = \mathbf{x}'\bar{\mathbf{x}}' = (2\mathbf{u}\mathbf{u}^{*'})(2\overline{\mathbf{u}\mathbf{u}^{*'}}) = 4\mathbf{u}(\mathbf{u}^{*'}\bar{\mathbf{u}}^{*'})\bar{\mathbf{u}} = 4r|\mathbf{u}'|^2 \tag{131}$$

and with the help of relation 131, we can turn equation (114) into

$$\frac{1}{2r^2}4r|\mathbf{u}'|^2 - \frac{\mu}{r} = -h. \tag{132}$$

Then, multiplying by $-r$ yields

$$\mu - 2|\mathbf{u}'|^2 = rh. \tag{133}$$

Now, substituting equation 133 into equation 130 yields

$$2r\mathbf{u}^{*''} + rh\mathbf{u}^* = 0 \tag{134}$$

and finally, by dividing by r gives us the quaternion differential equation we were looking for

$$2\mathbf{u}^{*''} + h\mathbf{u}^* = 0. \tag{135}$$

Theorem 3.3. By translating the Kepler problem to quaternion notation, and using the bilinear differential condition 103 and the fictitious time τ transformation 112, we have derived the differential equation:

$$2\mathbf{u}'' + h\mathbf{u} = 0, \tag{136}$$

which describes the motion of four uncoupled harmonic oscillators with common frequency $\omega := \sqrt{h/2}$. Note that this equation is linear in $\mathbf{u}$ and has no singularity at the origin. The equation looks very similar to equation (61) obtained by the Levi-Civita regularization, but whereas the regularization of Levi-Civita is limited to two dimensions, the regularization using quaternions can be used to describe three dimensional problems.

We can use Theorem 3.3 to derive the equations of motion describing the trajectory of the Kepler problem's orbital body in terms of the eccentric anomaly E. As described before, when no additional forces are present, conservation of angular momentum forces the trajectory of the orbital particle to be restricted to the plane spanned by its velocity vector and the separation vector between the orbital particle and the central body at the origin. Thus, the Kepler problem is reduced to two dimensions which allows us to use Levi-Civita's solution:

$$\mathbf{x} = x_0 + ix_1 = \mathbf{u^2}. \tag{137}$$

To find the equations of motion of the orbital particle, having found the transformation of the Kepler problem $\ddot{\mathbf{x}} + \mu\mathbf{x}r^{-3} = 0 \in C$, $r = |\mathbf{x}| = |\mathbf{u}|^2$ to the quaternion differential equation $2\mathbf{u}'' + h\mathbf{u} = 0 \in C$ with $dt = rd\tau$ and $h = -\frac{1}{2}|\dot{x}|^2 + \mu r^{-1} = \mathit{constant} > 0$, we will give the general solution of equation 137 in two dimensions:

$$\mathbf{u} = A\cos\omega\tau + iB\sin\omega\tau \in C, \tag{138}$$

with $\omega = \sqrt{h/2}$. Equation (139) parametrizes the origin-centered elliptic orbit of a planar harmonic oscillator in terms of τ. A and B are determined by the position and velocity vectors; the so called initial conditions of the system and for simplicity we assume them to be real, $A, B \in \Re$, which corresponds to the usage of a coordinate system that is aligned with the principle aces of the orbit and $\tau = 0$ corresponding to an apex of the ellipse.

We square equation (139) to obtain the final solution of this section:

$$\mathbf{x} = \mathbf{u}^2 = \frac{A^2 - B^2}{2} + \frac{A^2 + B^2}{2}\cos 2\omega\tau + iAB\sin 2\omega\tau \tag{139}$$

This is the solution of the elliptic trajectory of the orbital body of the Kepler problem, with the central body, the origin of the system, as one of its foci. This solution is analyzed in *Appendix 3: The resulting orbital motion from the Kepler problem.*

3.2 Solving the Kepler problem using spinors

We will begin this section with a summary of the relevant geometric algebra, which we will then apply to the Kepler problem in order to find the resulting equations of motion of the orbital body around the central body at the origin.

3.2.1 Geometric algebra

Geometric algebra is an associative algebra of a finite dimensional vector space over real scalars. The appeal of geometric algebra lies in the fact that elements and operations have a direct geometrical interpretation. The most important concept in geometric algebra is rule for vector multiplication, which together with vector space addition forms the geometric algebra. The geometric product is not limited to three dimensions, contrary to the cross product (in two dimensions the vector $\vec{a} \times \vec{b}$ needs a third dimension to point out of the plane spanned by $\vec{a}$ and $\vec{b}$ and in four dimensions the vector orthogonal to the vectors $\vec{a}$ and $\vec{b}$ is not unique).

Axiom 4. Let $\mathbb{X}$ be a finite dimensional vector space, then for every $\vec{a}$, $\vec{b}$, $\vec{c} \in \mathbb{X}$ there is a geometric product with the following properties:

$$\begin{aligned}
(\vec{a}\vec{b})\vec{c} &= \vec{a}(\vec{b}\vec{c}), \\
\vec{a}(\vec{b}+\vec{c}) &= \vec{a}\vec{b} + \vec{a}\vec{c}), \\
(\vec{a}+\vec{b})\vec{c} &= \vec{a}\vec{c} + \vec{b}\vec{c}, \\
\vec{a}\vec{a} &= \vec{a}^2 \geq 0 \in \Re,
\end{aligned}$$

and if $\vec{a}^2 \neq 0$, then:

$$\vec{a}^{-1} = \frac{\vec{a}}{\vec{a}^2}. \tag{140}$$

Thus, the geometric product is associative, and distributive over vector addition. Although the geometric product in general does not commute, it does commute with scalars. The geometric product $\vec{a}\vec{b}$ can be written as a sum of a symmetric and an antisymmetric part:

$$\vec{a}\vec{b} := \frac{1}{2}(\vec{a}\vec{b} + \vec{b}\vec{a}) + \frac{1}{2}(\vec{a}\vec{b} - \vec{b}\vec{a}) = \vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b}. \tag{141}$$

The symmetric part is given by the scalar product

$$\vec{a} \cdot \vec{b} = \frac{1}{2}(\vec{a}\vec{b} + \vec{b}\vec{a}) = \vec{b} \cdot \vec{a} \tag{142}$$

and the antisymmetric part is given by the so-called "wedge" product

$$\vec{a} \wedge \vec{b} = \frac{1}{2}(\vec{a}\vec{b} - \vec{b}\vec{a}) = -\vec{b} \wedge \vec{a}. \tag{143}$$

From this we can see that

$$\vec{a} \wedge \vec{a} = -\vec{a} \wedge \vec{a} = 0, \tag{144}$$

so the wedge product of two collinear vectors is 0 (because of associativity with scalar multiplication). Now, let $\vec{c} = \vec{a} + \vec{b}$, then

$$\vec{c}^2 = (\vec{a}+\vec{b})^2 = \vec{a}^2 + \vec{a}\vec{b} + \vec{b}\vec{a} + \vec{b}^2 = \vec{a}^2 + 2\vec{a}\cdot\vec{b} + \vec{b}^2, \tag{145}$$

$$\vec{a} \cdot \vec{b} = \frac{1}{2}(\vec{c}^2 - \vec{a}^2 - \vec{b}^2) = |\vec{a}||\vec{b}| \cos\theta, \tag{146}$$

where θ is the angle between two vectors $\vec{a}$ and $\vec{b}$. Furthermore, we have

$$(\vec{a} \wedge \vec{b})^2 = (\vec{a} \wedge \vec{b})(\vec{a} \wedge \vec{b}) = -(\vec{a} \wedge \vec{b})(\vec{b} \wedge \vec{a}) \tag{147}$$

$$= -(\vec{a}\vec{b} - \vec{a} \cdot \vec{b})(\vec{b}\vec{a} - \vec{b} \cdot \vec{a}) = -\left(\vec{a}\vec{b}\vec{b}\vec{a} - (\vec{a}\vec{b})(\vec{b} \cdot \vec{a}) - (\vec{a} \cdot \vec{b})(\vec{b}\vec{a}) + (\vec{a} \cdot \vec{b})^2\right) \tag{148}$$

$$= -\left(\vec{a}\vec{b}^2\vec{a} - (\vec{a} \cdot \vec{b})(\vec{a}\vec{b}) - (\vec{a} \cdot \vec{b})(\vec{b}\vec{a}) + (\vec{a} \cdot \vec{b})^2\right) \tag{149}$$

$$= -\left(\vec{a}\vec{a}\vec{b}^2 - (\vec{a} \cdot \vec{b})(\vec{a}\vec{b} + \vec{b}\vec{a}) + (\vec{a} \cdot \vec{b})^2\right) = -\left(\vec{a}^2\vec{b}^2 - 2(\vec{a} \cdot \vec{b})^2 + (\vec{a} \cdot \vec{b})^2\right) \tag{150}$$

$$= -(\vec{a}^2\vec{b}^2 - \vec{a}^2\vec{b}^2 \cos\theta^2) = -\vec{a}^2\vec{b}^2(1 - \cos^2\theta) = -\vec{a}^2\vec{b}^2 \sin^2\theta. \tag{151}$$

Thus, $(\vec{a} \wedge \vec{b})^2 \leq 0$ and $\vec{a} \wedge \vec{b}$ is proportional to $\sin\theta$. For two orthogonal vectors $\vec{a} \cdot \vec{b} = 0$, so $\vec{a}\vec{b} = \vec{a} \wedge \vec{b}$. If $\vec{\sigma}_i$ and $\vec{\sigma}_j$ are two orthonormal unit vectors then

$$(\vec{\sigma}_i\vec{\sigma}_j)^2 = (\vec{\sigma}_i \wedge \vec{\sigma}_j)^2 = -\vec{\sigma}_i^2\vec{\sigma}_j^2 \sin^2\theta = -\sin^2\theta = -1. \tag{152}$$

So far, we have seen some of the properties of the wedge product, but we have not defined it yet.

Let $\vec{\sigma}_1, \vec{\sigma}_2, ..., \vec{\sigma}_n$ be an orthonormal basis for the set of linearly independent vectors $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$, so we can write

$$\vec{a}_i = \sum_j \alpha_{ij}\vec{\sigma}_j. \tag{153}$$

Here, α is the matrix of coefficients of the vectors $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$:

$$\alpha = \begin{pmatrix} \vec{\alpha}_{1,1} & \vec{\alpha}_{1,2} & \cdots & \vec{\alpha}_{1,n} \\ \vec{\alpha}_{2,1} & \vec{\alpha}_{2,2} & \cdots & \vec{\alpha}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{\alpha}_{n,1} & \vec{\alpha}_{n,2} & \cdots & \vec{\alpha}_{n,n} \end{pmatrix}.$$

We define the wedge product as:

$$\vec{a}_1 \wedge ... \wedge \vec{a}_n := \det(\alpha)\vec{\sigma}_1...\vec{\sigma}_n. \tag{154}$$

Thus

$$\vec{a}_1 \wedge ... \wedge (\vec{a}_j + \vec{b}_j) \wedge ... \wedge \vec{a}_n = \vec{a}_1 \wedge ... \wedge \vec{a}_j \wedge ... \wedge \vec{a}_n + \vec{a}_1 \wedge ... \wedge \vec{b}_j \wedge ... \wedge \vec{a}_n \tag{155}$$

and

$$\vec{a}_1 \wedge ... \wedge \vec{a}_j \wedge \vec{a}_{j+1} \wedge ... \wedge a_n = -\vec{a}_1 \wedge ... \wedge \vec{a}_{j+1} \wedge \vec{a}_j \wedge ... \wedge \vec{a}_n. \tag{156}$$

The wedge product of two vectors $\vec{a} \wedge \vec{b}$ signifies a bivector, determined by the vectors $\vec{a}$ and $\vec{b}$. Bivectors have a clear geometrical interpretation. A regular vector has an orientation an magnitude. It can be interpreted as a directed line segment, where the orientation represents a directed line and the magnitude of the vector signifies the length of the segment of this line that the (regular) vector consists of. Similarly, we can interpret the bivector as an oriented plane. The direction of the bivector represents the orientation of the plane and the magnitude of the bivector signifies the area of the plane. The plane the bivector represents is oriented in such a way that it represents the plane in which the vectors $\vec{a}$ and $\vec{b}$ lie. The area of this plane is given by the area of a parallelogram with sides $\vec{a}$ and $\vec{b}$. The bivector should be regarded as an seperate entity, not just the product $\vec{a} \wedge \vec{b}$ of two vectors $\vec{a}$ and $\vec{b}$. From here on we will denote bivectors, as well as higher dimensional multivectors with capital letters (multivectors will be specified later on).

The geometric product in general does not commute. However, we can use the "degree of commutativity" of the vectors $\vec{a}$ and $\vec{b}$ under the geometric product as a measure of the angle between the two vectors. If the two vectors $\vec{a}$ and $\vec{b}$ are collinear, we get:

$$\vec{a}\vec{b} = \vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b} = \vec{a} \cdot \vec{b} + 0 = \vec{b} \cdot \vec{a} - 0 = \vec{b} \cdot \vec{a} + \vec{b} \wedge \vec{a} = \vec{b}\vec{a}. \tag{157}$$

We can understand that the wedge product of two collinear vectors is equal to zero by referring back to the geometric interpretation of the bivector. The "parallelogram" with sides $\vec{a}$ and $\vec{b}$, where $\vec{a}$ and $\vec{b}$ are collinear, is a line segment. Thus, it is a "plane segment" with zero area and so their wedge product equals zero. Thus, we conclude that collinear vectors commute under the geometric product: $\vec{a}\vec{b} = \vec{b}\vec{a}$, because the antisymmetric part of the geometric product $\vec{a}\vec{b}$ vanishes. For orthogonal vectors on the other hand we have

$$\vec{a}\vec{b} = \vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b} = 0 + \vec{a} \wedge \vec{b} = 0 - \vec{b} \wedge \vec{a} = -\vec{b}\vec{a}. \tag{158}$$

As we have seen before, the geometric product of two orthogonal vectors is equal to the wedge product of these vectors, because the scalar product of two orthogonal vectors vanishes. Thus, two orthogonal vectors $\vec{a}$ and $\vec{b}$ anti-commute under the geometric product. With the *degree of commutativity* we mean how close the geometric product $\vec{a}\vec{b}$ of two vectors $\vec{a}$ and $\vec{b}$ gets to obeying either relation (157) or (158). The degree of commutativity describes how the geometric product $\vec{a}\vec{b}$ of two vectors $\vec{a}$ and $\vec{b}$ lies between the extremes of orthogonal or collinear vectors. If relation (157) *almost* holds, it tells us that the angle between the two vectors $\vec{a}$ and $\vec{b}$ is very small and that they are almost collinear. Similarly, if relation (158) almost holds, we know the vectors $\vec{a}$ and $\vec{b}$ are almost orthogonal.

For constructing a geometric product we choose a right-handed frame of orthonormal unit vectors $\vec{\sigma}_1, \vec{\sigma}_2, \vec{\sigma}_3$, for which the geometric product is defined by the relationship

$$\vec{\sigma}_i\vec{\sigma}_j + \vec{\sigma}_j\vec{\sigma}_i := 2\delta_{ij}. \tag{159}$$

Because the vectors the basis vectors are orthonormal, we have

$$\begin{aligned}
\vec{\sigma}_1^2 = \vec{\sigma}_1 \cdot \vec{\sigma}_1 + \vec{\sigma}_1 \wedge \vec{\sigma}_1 = \vec{\sigma}_1 \cdot \vec{\sigma}_1 = 1 = \vec{\sigma}_2^2 = \vec{\sigma}_3^2 \qquad (160)\\
\vec{\sigma}_1\vec{\sigma}_2 = \vec{\sigma}_1 \cdot \vec{\sigma}_2 + \vec{\sigma}_1 \wedge \vec{\sigma}_2 = \vec{\sigma}_1 \wedge \vec{\sigma}_2 = -\vec{\sigma}_2\vec{\sigma}_1 \\
\vec{\sigma}_2\vec{\sigma}_3 = \vec{\sigma}_2 \cdot \vec{\sigma}_3 + \vec{\sigma}_2 \wedge \vec{\sigma}_3 = \vec{\sigma}_2 \wedge \vec{\sigma}_3 = -\vec{\sigma}_3\vec{\sigma}_2 \\
\vec{\sigma}_3\vec{\sigma}_1 = \vec{\sigma}_3 \cdot \vec{\sigma}_1 + \vec{\sigma}_3 \wedge \vec{\sigma}_1 = \vec{\sigma}_3 \wedge \vec{\sigma}_1 = -\vec{\sigma}_1\vec{\sigma}_3.
\end{aligned}$$

(161)

Consider now how a vector $\vec{a}$ can be expressed as a linear combination of basis vectors

$$\vec{a} = \sum_i a_1 \vec{\sigma}_i. \tag{162}$$

Similarly, a bivector B can be expressed as a linear combination of the basis vectors $\sigma_i \wedge \sigma_j$

$$B = \sum_i \sum_j \frac{1}{2} \; B_{ij} \vec{\sigma}_i \wedge \vec{\sigma}_j. \tag{163}$$

The scalars $B_{ij} = -B_{ji}$ (since $\vec{\sigma}_i \wedge \vec{\sigma}_j = -\vec{\sigma}_j \wedge \vec{\sigma}_i$) are called components of B with respect to the basis $\vec{\sigma}_i \wedge \vec{\sigma}_j$. From (160) we can see that bivectors compose a three-dimensional linear space with basis vectors $\vec{\sigma}_1\vec{\sigma}_2$, $\vec{\sigma}_2\vec{\sigma}_3$ and $\vec{\sigma}_3\vec{\sigma}_1$. We can write the geometric product $\vec{a}\vec{b}$ in terms of the scalar product $\vec{a} \cdot \vec{b}$ and the bivector B with components $B_{ij} = 2a_i b_i$. The bivector B with components $B_{ij} = 2a_i b_i$ is given by:

$$B = \vec{a} \wedge \vec{b} = \sum_i \sum_j \frac{1}{2} \; B_{ij} \vec{\sigma}_i \wedge \vec{\sigma}_j = a_2 b_3 \vec{\sigma}_2 \wedge \vec{\sigma}_3 + a_3 b_2 \sigma_3 \wedge \vec{\sigma}_2 ... \tag{164}$$

$$... + a_3 b_1 \vec{\sigma}_3 \wedge \vec{\sigma}_1 + a_1 b_3 \vec{\sigma}_1 \wedge \vec{\sigma}_3 + a_1 b_2 \vec{\sigma}_1 \wedge \vec{\sigma}_2 + a_2 b_1 \vec{\sigma}_2 \wedge \vec{\sigma}_1$$

$$= (a_2 b_3 - a_3 b_2) \vec{\sigma}_2 \wedge \vec{\sigma}_3 + (a_3 b_1 - a_1 b_3) \vec{\sigma}_3 \wedge \vec{\sigma}_1 + (a_1 b_2 - a_2 b_1) \vec{\sigma}_1 \wedge \vec{\sigma}_2 \tag{165}$$

$$= (a_2 b_3 - a_3 b_2) \vec{\sigma}_2 \vec{\sigma}_3 + (a_3 b_1 - a_1 b_3) \vec{\sigma}_3 \vec{\sigma}_1 + (a_1 b_2 - a_2 b_1) \vec{\sigma}_1 \vec{\sigma}_2. \tag{166}$$

Here, I have not written out the terms $a_i b_i \vec{\sigma}_i \wedge \vec{\sigma}_i$, since $\vec{\sigma}_i \wedge \vec{\sigma}_i = 0$. We can see the resemblance with the coefficients of the cross product $\vec{a} \times \vec{b}$.

We have introduced the bivectors $\vec{\sigma}_i \wedge \vec{\sigma}_j$ bij multiplying together the $\vec{\sigma}_i$ under the wedge product. There is one more concept we can introduce by multiplying $\vec{\sigma}_i$, the *pseudo scalar* or *trivector* I:

$$I = \vec{\sigma}_1 \vec{\sigma}_2 \vec{\sigma}_3 = \vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3. \tag{167}$$

We can give a geometric interpretation to the pseudo scalar just as we did with the bivector. As the bivector $\vec{\sigma}_1 \wedge \vec{\sigma}_2$ represented an oriented plane with unit area, the trivector $\vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3$ represents an oriented three-space segment with unit volume. As with the bivector, the pseudo scalar should not be considered just as the product $\vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3$, but as an entity of its own, with a geometric interpretation. In three dimensions, the trivector is as "high" as we can go, since the geometric product of four orthonormal basis vectors will contain at least one of the basis vectors twice. In that case the geometric product of four orthonormal basis vectors will reduce to a bivector by the elimination of the doubly occurring basis vector ($\vec{\sigma}_i \vec{\sigma}_i = 1$). Given $\vec{a} = \sum_{i=1}^{3} a_i \vec{\sigma}_i$ and $\vec{b} = \sum_{i=1}^{3} b_i \vec{\sigma}_i$, the pseudo scalar I allows us to write the the bivector B with components $B_{ij} = 2a_i b_i$ as

$$\vec{a} \wedge \vec{b} = I \vec{a} \times \vec{b}. \tag{168}$$

Proof: See proof 6.9 in Appendix 1: Proofs. Thus, given $\vec{a} = \Sigma_{i=1}^{3} a_i\vec{\sigma}_i$ and $\vec{b} = \Sigma_{i=1}^{3} b_i\vec{\sigma}_i$ the geometric product $\vec{a}\vec{b}$ is given by

$$\vec{a}\vec{b} = a_1b_1 + a_2b_2 + a_3b_3 + (a_2b_3 - a_3b_2)\vec{\sigma}_2 \wedge \vec{\sigma}_3 - (a_1b_3 - a_3b_1)\vec{\sigma}_3 \wedge \vec{\sigma}_1 + (a_1b_2 - a_2b_1)\vec{\sigma}_1 \wedge \vec{\sigma}_2, \tag{169}$$

where we have used

$$(\vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3)\vec{\sigma}_3 = \vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge (\vec{\sigma}_3\vec{\sigma}_3) = \vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge (\vec{\sigma}_3 \cdot \vec{\sigma}_3 + \vec{\sigma}_3 \times \vec{\sigma}_3)$$

$$= \vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge (1 + 0) = \vec{\sigma}_1 \wedge \vec{\sigma}_2. \tag{170}$$

Similarly,

$$(\vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3)\vec{\sigma}_2 = -(\vec{\sigma}_1 \wedge \vec{\sigma}_3 \wedge \vec{\sigma}_2)\vec{\sigma}_2 = -\vec{\sigma}_1 \wedge \vec{\sigma}_3, \tag{171}$$

$$(\vec{\sigma}_1 \wedge \vec{\sigma}_2 \wedge \vec{\sigma}_3)\vec{\sigma}_1 = (\vec{\sigma}_2 \wedge \vec{\sigma}_3 \wedge \vec{\sigma}_1)\vec{\sigma}_1 = \vec{\sigma}_2 \wedge \vec{\sigma}_3. \tag{172}$$

We can use the unit trivector to rewrite the geometric product as the sum of the scalar product and the product of the unit trivector and the vector cross-product

$$\vec{a}\vec{b} = \vec{a} \cdot \vec{b} + I\vec{a} \times \vec{b}. \tag{173}$$

Moreover, we can extract the scalar product and the vector cross-product from the geometric product. Namely, by using the properties that the scalar product is symmetric and that the vector cross-product is anti-symmetric:

$$\frac{1}{2}(\vec{a}\vec{b} + \vec{b}\vec{a}) = \frac{1}{2}(\vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b} + \vec{b} \cdot \vec{a} + \vec{b} \wedge \vec{a} = \frac{1}{2}(2\vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b} - \vec{a} \wedge \vec{b}) = \vec{a} \cdot \vec{b}, \tag{174}$$

$$\frac{1}{2I}(\vec{a}\vec{b} - \vec{b}\vec{a}) = \frac{1}{2I}(\vec{a} \cdot \vec{b} + \vec{a} \wedge \vec{b} - \vec{b} \cdot \vec{a} - \vec{b} \wedge \vec{a}) = \frac{1}{2I}(2\vec{a} \wedge \vec{b}) = \frac{1}{2I}(2I\vec{a} \times \vec{b}) = \vec{a} \times \vec{b}. \tag{175}$$

Equation (173) consists of both a scalar an a bivector. A multivector is an arbitrary element of gemetric algebra. In three dimensions we can state that any multivector can be broken down into the sum of four quantities: a scalar, a vector, a bivector and a trivector. Thus, every element of the geometric algebra of physical space can be written as a linear combination of the basis elements $\{1, \vec{\sigma}_i, \vec{\sigma}_i \wedge \vec{\sigma}_j, \vec{\sigma}_i \wedge \vec{\sigma}_j \wedge \vec{\sigma}_k\}$, consisting of the vectors $\vec{\sigma}_i$ and their products. Thus, a multivector M can be written as the sum of a scalar M_S, a vector M_V, a bivector M_B and a trivector M_P. Alternatively, using the k-vector nomenclature, the scalar part is called 0-vector, the vector part 1-vector, the bivector part 2-vector and the trivector part 3-vector. Thus, an arbitrary multivector (in three dimensions) $M = M_S + M_V + M_B + M_P$ can be expressed as the sum of a 0-vector, a 1-vector, a 2-vector and a 3-vector. The *grade projector* $\langle . \rangle_k$ can be used to select the k-vector part, $k = 0, 1, 2, 3$, of an arbitrary multivector. Choose $\langle A \rangle_k$ to be the k-vector part of the multivector A. The scalar projector $\langle A \rangle_0$ is often simply written as $\langle A \rangle$. A k-vector is said to be even if k is even, and odd if k is odd. Thus, scalars and bivectors are even, whereas vectors and trivectors are odd. An odd multivector M_O

consists of a vector and a trivector. Similarly, an even multivector M_E consists of the sum of a scalar and a bivector, and is also called a *spinor*. Because we discussed the importance of the geometric interpretation of geometric algebra, it is natural to wonder what is means to add scalars, vectors, bivectors, and/or trivectors into a multivector. Almost every combination of spinors, vectors, bivectors and/or trivectors has turned out to have a significant physical interpretation, as for example electric fields are best represented by a multivector consisting of a vector and a bivector [4]. Spinors turn out to be very useful for representing rotation-dilatations in three dimensions and will be used in the following sections to regularize the kepler problem. Since we will use spinors to reguralize the Kepler problem, we will focus on spinors in the following section. We will use capital letters to denote bi-, tri- and multivectors in general. Specifically, we will use bold-face capital letters to denote spinors. In three dimensions the wedge product of two bivectors vanishes [8]. The geometric product of two bivectors is calculated as follows. Let $U, V \in M_B$, $U = u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3$ and $V = v_1\vec{\sigma}_1\vec{\sigma}_2 + v_2\vec{\sigma}_1\vec{\sigma}_3 + v_3\vec{\sigma}_2\vec{\sigma}_3$ then

$$UV = (u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3)(v_1\vec{\sigma}_1\vec{\sigma}_2 + v_2\vec{\sigma}_1\vec{\sigma}_3 + v_3\vec{\sigma}_2\vec{\sigma}_3) \tag{176}$$

$$= -u_1v_1 - u_2v_2 - u_3v_3 + (u_2v_3 - u_3v_2)\vec{\sigma}_1\vec{\sigma}_2 - (u_1v_3 - u_3v_1)\vec{\sigma}_1\vec{\sigma}_3 + (u_1v_2 - u_2v_1)\vec{\sigma}_2\vec{\sigma}_3 \tag{177}$$

$$= U \cdot V + U \times V. \tag{178}$$

The symmetric part is given by

$$U \cdot V = -u_1v_1 - u_2v_2 - u_3v_3 = -v_1u_1 - v_2u_2 - v_3u_3 = V \cdot U \tag{179}$$

and the anti-symmetric part is given by

$$U \times V = (u_2v_3 - u_3v_2)\vec{\sigma}_1\vec{\sigma}_2 - (u_1v_3 - u_3v_1)\vec{\sigma}_1\vec{\sigma}_3 + (u_1v_2 - u_2v_1)\vec{\sigma}_2\vec{\sigma}_3 \tag{180}$$

$$= -(v_2u_3 - v_3u_2)\vec{\sigma}_1\vec{\sigma}_2 + (v_1u_3 - v_3u_1)\vec{\sigma}_1\vec{\sigma}_3 - (v_1u_2 - v_2u_1)\vec{\sigma}_2\vec{\sigma}_3) = -V \times U. \tag{181}$$

The geometric product of two spinors $\mathbf{U}, \mathbf{V} \in M_E$ now becomes

$$\mathbf{UV} = (u_0 + u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3)(v_0 + v_1\vec{\sigma}_1\vec{\sigma}_2 + v_2\vec{\sigma}_1\vec{\sigma}_3 + v_3\vec{\sigma}_2\vec{\sigma}_3) \tag{182}$$

$$= (\langle\mathbf{U}\rangle_\mathbf{0} + \langle\mathbf{U}\rangle_\mathbf{2})(\langle\mathbf{V}\rangle_\mathbf{0} + \langle\mathbf{V}\rangle_\mathbf{2}) \tag{183}$$

$$= \langle\mathbf{V}\rangle_\mathbf{0}\langle\mathbf{U}\rangle_\mathbf{0} + \langle\mathbf{V}\rangle_\mathbf{0}\langle\mathbf{U}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{0}\langle\mathbf{V}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{2} \cdot \langle\mathbf{V}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{2} \times \langle\mathbf{V}\rangle_\mathbf{2}. \tag{184}$$

Using equations (179) and (180), we have

$$= u_0v_0 + u_0(v_1\vec{\sigma}_1\vec{\sigma}_2 + v_2\vec{\sigma}_1\vec{\sigma}_3 + v_3\vec{\sigma}_2\vec{\sigma}_3) + v_0(u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3)... \tag{185}$$

$$...-u_1v_1-u_2v_2-u_3v_3+(u_2v_3-u_3v_2)\vec{\sigma}_1\vec{\sigma}_2-(u_1v_3-u_3v_1)\vec{\sigma}_1\vec{\sigma}_3+(u_1v_2-u_2v_1)\vec{\sigma}_2\vec{\sigma}_3 \quad (186)$$

$$=(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+(u_0v_1+u_1v_0+u_2v_3-u_3v_2)\vec{\sigma}_1\vec{\sigma}_2... \quad (187)$$

$$...+(u_0v_2+u_2v_0-u_1v_3+u_3v_1)\vec{\sigma}_1\vec{\sigma}_3+(u_0v_3+u_3v_0+u_1v_2-u_2v_1)\vec{\sigma}_2\vec{\sigma}_3.$$

The geometric product of the spinor shown above is associative. Let $\mathbf{U},\mathbf{V},\mathbf{W}\in M_E$, then

$$(\mathbf{UV})\mathbf{W}=\mathbf{U}(\mathbf{VW}). \quad (188)$$

Proof: See proof 6.10 in Appendix 1: Proofs.

Whenever an equation contains both a geometric product and either a scalar product, a vector cross-product or all three, the general convention is that the scalar and/or vector products take precedence over the geometric product. We will also define the reversion $A^\dagger$ of a multivector. The reversion $A^\dagger$ is obtained by an interchange of the order of the vectors in a geometric product. Thus, scalars and vectors remain unchanged under reversion. However, bivectors and trivectors change their sign under reversion (since $\vec{\sigma}_i\vec{\sigma}_j=-\vec{\sigma}_j\vec{\sigma}_i$). A spinor $\mathbf{S}\in M_E$ changes under the reversion as follows:

$$\mathbf{S}^\dagger:=(\langle\mathbf{S}\rangle_\mathbf{0}+\langle\mathbf{S}\rangle_\mathbf{2})^\dagger=\langle\mathbf{S}\rangle_\mathbf{0}-\langle\mathbf{S}\rangle_\mathbf{2}, \quad (189)$$

because scalars remain unchanged under reversion and bivectors change their sign. For the product of two spinors $\mathbf{A},\mathbf{B}\in M_E$

$$(\mathbf{AB})^\dagger=\mathbf{B}^\dagger\mathbf{A}^\dagger. \quad (190)$$

Proof: see proof 6.11 in Appendix 1: Proofs. The *modulus* of a spinor $\mathbf{S}\in M_E$ $\mathbf{S}=s_0+s_1\vec{\sigma}_1\vec{\sigma}_2+s_2\vec{\sigma}_1\vec{\sigma}_3+s_3\vec{\sigma}_2\vec{\sigma}_3$ is given by

$$|\mathbf{S}|^\mathbf{2}:=\mathbf{SS}^\dagger=s_0^2-s_1^2-s_2^2-s_3^2. \quad (191)$$

Proof: see proof 6.12 in Appendix 1: Proofs. We also introduce the *inverse* of a spinor $\mathbf{S}\in M_E$

$$\mathbf{S}^{-\mathbf{1}}:=\frac{\mathbf{S}^\dagger}{|\mathbf{S}|^\mathbf{2}}=\frac{\mathbf{S}^\dagger}{\mathbf{SS}^\dagger} \quad (192)$$

and, as expected from the inverse, we see that

$$\mathbf{S}^{-\mathbf{1}}\mathbf{S}=\frac{\mathbf{S}^\dagger\mathbf{S}}{\mathbf{SS}^\dagger}=\frac{\mathbf{SS}^\dagger}{\mathbf{SS}^\dagger}=\mathbf{SS}^{-\mathbf{1}}=1. \quad (193)$$

Rotation are written in the following form in geometric algebra

$$\vec{a}\mapsto\mathbf{R}(\vec{a})=\mathbf{R}\vec{a}\mathbf{R}^\dagger \quad (194)$$

The following normalization condition is satisfied by any even multivector $\mathbf{R}$

$$\mathbf{R}\mathbf{R}^\dagger = 1. \tag{195}$$

Any normalized even multivector represents a rotation. For an arbitrary, not necessarily normalized, even multivector the following holds:

$$\alpha = \mathbf{U}\mathbf{U}^\dagger \geq 0. \tag{196}$$

This makes $U = \sqrt{\alpha}\mathbf{R}$ a multiple of a rotor $\mathbf{R}$, which allows us to write

$$\mathbf{U}\vec{a}\mathbf{U}^\dagger = \alpha\mathbf{R}\vec{a}\mathbf{R}^\dagger. \tag{197}$$

$\mathbf{U}$ as such represents a rotation-dilatation of three-space. For any even multivector $\mathbf{U}$ and any vector $\vec{a}$ we have:

$$\mathbf{U}\vec{a}\mathbf{U}^\dagger = \langle\mathbf{U}\vec{a}\mathbf{U}^\dagger\rangle_1. \tag{198}$$

Having two vectors $\vec{a}$ and $\vec{b}$, a rotation can be specified by the rotation in the plane spanned by $\vec{a}$ and $\vec{b}$, $\langle\vec{a}\vec{b}\rangle_2$, mapping $\vec{a}$ to $\vec{b}$. This rotation is given by the rotor

$$\mathbf{R} = \frac{1 + \vec{b}\vec{a}}{|\vec{a} + \vec{b}|} = \frac{1 + \vec{b}\vec{a}}{\sqrt{2(1 + \vec{a} \cdot \vec{b})}}. \tag{199}$$

Using the normalization condition, for $|\vec{a} + \vec{b}|$ we have

$$|\vec{a} + \vec{b}| = \sqrt{(a_1 + b_1)^2 + (a_2 + b_2)^2 + (a_3 + b_3)^2} \tag{200}$$

$$= \sqrt{\sum_{i=1}^{3} a_i^2 + \sum_{i=1}^{3} b_i^2 + 2\sum_{i=1}^{3} a_i b_i} = \sqrt{1 + 1 + 2\vec{a} \cdot \vec{b}}. \tag{201}$$

We can also write the rotor in terms of its rotation axis, given by a unit vector $\vec{n}$ and a rotation angle ϕ

$$\mathbf{R} = e^{-I\vec{n}\phi/2} \tag{202}$$

The definition of this exponential is given by the usual power series expansion

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \tag{203}$$

and, for $AB = BA$, it satisfies:

$$e^{A+B} = e^A e^B. \tag{204}$$

Additionally the following properties hold:

$$e^A B = B e^A \quad \text{if } AB = BA \tag{205}$$

$$e^A B = B e^{-A} \quad \text{if } AB = -BA \tag{206}$$

We also need to consider differential calculus on the geometric product. The multivector derivative provides differential calculus for an arbitrary multivector. Choose a smooth multivector-valued function $F(X)$ of the multivector X. For an arbitrary multivector A, the directional derivative in the direction of A is given by

$$A \cdot \partial_x F(X) = \left.\frac{dF(X + \tau P - x(A))}{d\tau}\right|_{\tau=0}. \tag{207}$$

Here, $P_x(A)$ projects A onto the grades contained in X. Let e_J, $J = 1, ..., 8$ be a basis of the geometric algebra and e^J its dual basis, such that $e^J \cdot e^K = \delta_J^K$. The multivector derivative is then given by

$$\partial_x = \sum_J e^J e_J \cdot \delta_x. \tag{208}$$

For a vector argument $\mathbf{x}$, the derivative reduces to the vector derivative. As you can see, the scalar product $A \cdot \delta_x$ gives the directional derivative on A. Additionally, note that it inherits the algebraic properties of the argument X; δ_x consists of the same grades as X. The Leibniz rule holds for both the directional derivative and the multivector derivative, both linear operators, since:

$$A \cdot \partial_x (F(X)G(X)) = (A \cdot \partial_x F(X))G(X) + F(X)(A \cdot \partial_x G(X)), \tag{209}$$

$$\partial_x (F(X)G(X)) = \dot{\partial}_x \dot{F}(X)G(X) + \dot{\partial} F(X)\dot{G}(X). \tag{210}$$

Where $(\dot{})$ denotes a function to be differentiated. The directional derivative satisfies the chain rule according to

$$A \cdot \partial_x F(G(X)) = (A \cdot \delta_x G(X)) \cdot \partial_x F(G) \tag{211}$$

and additionally

$$\partial_x \langle XA \rangle = \partial_x \langle AX \rangle = P_x(A), \tag{212}$$

which means that:

$$\partial_x \langle X^\dagger A \rangle = \partial_x \langle AX^\dagger \rangle = P_x(A^\dagger). \tag{213}$$

3.2.2 Solving the Kepler problem using spinors

We will now use spinors to represent the position of the orbital body around the central body not by a position vector, but by a rotation-dilation vector operator, a spinor, with reference to a fixed position vector. The position vector $\vec{x}$ is then given by the rotation dilation transformation of the reference vector by the spinor. As explained before, the rotation dilatation of the reference vector can be represented by an even multivector $\mathbf{U}$ in the form:

$$\vec{x} = \frac{1}{2}\mathbf{U}\vec{\sigma}_3\mathbf{U}^\dagger, \tag{214}$$

where $\vec{\sigma}_3$ is a fixed reference vector. The factor $\frac{1}{2}$ was inherited from earlier application, other than tradition there is no important reason to preserve it. It implies the normalization:

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = 2r = 2|\vec{x}|. \tag{215}$$

However, for a given position vector $\vec{x}$, the spinor that transforms the fixed reference vector $\vec{\sigma}_3$ into the position vector $\vec{x}$ is not unique:

$$\mathbf{U} \mapsto \mathbf{U}e^{-I_3\alpha/2}. \tag{216}$$

This mapping, for an arbitrary α does not transform the resulting position vector $\vec{x}$ the reference vector $\vec{\sigma}_3$ is mapped into, because it adds an additional exponential factor that corresponds to a rotation around the reference vector $\vec{\sigma}_3$ itself, which leaves the reference vector $\vec{\sigma}_3$ unchanged. With reference to equation (214), we can also write this mapping (216) that preserves the resulting position vector $\vec{x}$ as:

$$\mathbf{U} \mapsto \frac{\mathbf{U}}{\sqrt{2r}}e^{-I_3\alpha/2}\frac{\mathbf{U}^\dagger}{\sqrt{2r}}\mathbf{U} = e^{-I\vec{x}\alpha/2r}\mathbf{U}. \tag{217}$$

Now it represents a rotation of the resulting position vector $\vec{x}$ around itself, after being transformed into the position vector $\vec{x}$ from the reference vector $\vec{\sigma}_3$. Independent of whether the reference vector $\vec{\sigma}_3$ is first rotated around itself and then transformed into the position vector $\vec{x}$ or the other way around, this rotation of the vector around itself creates the need for introducing a fourth dimension when using a position spinor representation in three dimensions. Inversely, the position spinors corresponding to a position vector $\vec{x}$ can be found from

$$\mathbf{U} = \frac{r + \vec{x}\vec{\sigma}_3}{\sqrt{r+z}}e^{-I_3\alpha/2}, \tag{218}$$

for any arbitray α. The spinor U can be represented as $\mathbf{U} = u_0 + I\vec{u}$ with $\vec{u} = \Sigma_{k=1}^3 \vec{u}_k\vec{u}_k$, which decomposes transformation (214) into

$$\begin{aligned} \vec{x} &= u_1u_3 - u_0u_2 \\ \vec{y} &= u_1u_0 + u_2u_3 \\ \vec{z} &= \tfrac{1}{2}(u_0^2 - u_1^2 - u_2^2 + u_3^2) \end{aligned}$$

In order to obtain the equation of motion for $\mathbf{U}$, we need to derive the time derivatives of $\mathbf{U}$ by differentiating equation (214) with respect to time:

$$\dot{\vec{x}} = \frac{1}{2}\dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger + \frac{1}{2}\mathbf{U}\vec{\sigma}_3\dot{\mathbf{U}}^\dagger = \langle\dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger\rangle_1. \tag{219}$$

However, we cannot yet solve this for $\dot{\mathbf{U}}$, because we cannot determine the time derivative of α . Therefore, we will first impose a constraint on α. Referring to equation (216), we make the substitution

$$\dot{\mathbf{U}} \mapsto \dot{\mathbf{U}}e^{-I_3\alpha/2} - \frac{\dot{\alpha}}{2}\mathbf{U}I_3e^{-I_3\alpha/2}, \tag{220}$$

which allows us to write:

$$\dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger \mapsto \dot{\mathbf{U}}\sigma_3\mathbf{U}^\dagger - \dot{\alpha}rI. \tag{221}$$

Thus, we have introduced a pure trivector contribution into the product $\dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger$, which we can use to eliminate the trivector component completely by requiring

$$\langle\dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger\rangle_3 = 0 \tag{222}$$

This means that $\dot{\mathbf{U}}$ is chosen in such a way that it does not contain a component of motion in the direction of the degree of freedom, the rotation around itself. Requirement (222) thus requires that there does occur any rotation around the position vector $\vec{x}$ itself. By using the requirement (222), we can solve equation (219) to obtain:

$$\dot{\vec{x}} = \dot{\mathbf{U}}\vec{\sigma}_3\mathbf{U}^\dagger \tag{223}$$

and

$$\dot{\mathbf{U}} = \dot{\vec{x}}\mathbf{U}^{\dagger^{-1}}\vec{\sigma}_3 = \dot{\vec{x}}\frac{\mathbf{U}}{2r}\vec{\sigma}_3. \tag{224}$$

Now we introduce the *fictitious time* τ according to

$$dt = 2rd\tau, \tag{225}$$

such that, using (221), the derivative of U with respect to τ is given by

$$\mathbf{U}' = \frac{\mathbf{dU}}{d\tau} = \frac{\mathbf{dU}}{dt}\frac{dt}{d\tau} = 2r\dot{\mathbf{U}} = \dot{\vec{x}}\mathbf{U}\sigma_3. \tag{226}$$

Here $(\ )'$ signifies the derivative with respect to the fictitious time τ. A specific trajectory of the orbital body is specified by the initial position vector $\vec{x}$ and the initial velocity vector $\dot{\vec{x}}$, which can be converted into initial conditions for the spinor $\mathbf{U}$ by using equations (218) and (223). The second derivative of $\mathbf{U}$ is given by

$$\frac{d}{d\tau}\mathbf{U}' = \frac{d}{d\tau}\left(\dot{\vec{x}}\mathbf{U}\vec{\sigma}_3\right) = \left(\frac{d}{d\tau}\dot{\vec{x}}\right)\mathbf{U}\vec{\sigma}_3 + \dot{\vec{x}}\mathbf{U}''\vec{\sigma}_3, \tag{227}$$

$$\mathbf{U}'' = \frac{d\dot{\vec{x}}}{dt}\frac{dt}{d\tau}\mathbf{U}\vec{\sigma}_3 + \dot{\vec{x}}(\dot{\vec{x}}\mathbf{U}\vec{\sigma}_3)\vec{\sigma}_3. \tag{228}$$

By using identities (215) and (159) we can rewrite equation (228) as

$$\mathbf{U}'' = 2r\ddot{\vec{x}}\mathbf{U}\sigma_3\frac{\mathbf{U}^\dagger}{2r}\mathbf{U} + \dot{\vec{x}}^2\mathbf{U} \tag{229}$$

and using equation (214) we obtain

$$\mathbf{U}'' = 2\left(\ddot{\vec{x}}\vec{x} + \frac{1}{2}\dot{\vec{x}}^2\right)\mathbf{U}. \tag{230}$$

We now introduce the Kepler problem:

$$\ddot{\vec{x}} = -\frac{\vec{x}}{r^3}. \tag{231}$$

Which we substitute into identity (230) to obtain:

$$\mathbf{U}'' = 2\left((-\frac{\vec{x}}{r^3})\vec{x} + \frac{1}{2}\dot{\vec{x}}^2\right)\mathbf{U}, \tag{232}$$

$$= 2\left((-\frac{\vec{x}^2}{r^3}) + \frac{1}{2}\dot{\vec{x}}^2\right)\mathbf{U}, \tag{233}$$

$$= 2\left((-\frac{\vec{r}^2}{r^3}) + \frac{1}{2}\dot{\vec{x}}^2\right)\mathbf{U}, \tag{234}$$

$$= 2\left(\frac{1}{2}\dot{\vec{x}}^2 - \frac{1}{r}\right)\mathbf{U} = 2H\mathbf{U}, \tag{235}$$

where H corresponds to the Hamiltonian energy $H = \frac{1}{2}\dot{\vec{x}}^2 - \frac{1}{r}$, which consists of the sum of both the kinetic and the potential energy.

Theorem 3.4. By translating the Kepler problem to spinor notation, and using the trivector-contribution eliminating requirement (225) and the fictitious time τ transformation (222), we have derived the differential equation:

$$\mathbf{U}'' = 2H\mathbf{U}, \tag{236}$$

which describes the motion of four uncoupled harmonic oscillators with common frequency $\omega := \sqrt{-2H}$ for $H < 0$

This is the final result of this section. Using the same procedure as in section (3.1.2) to derive the equation of motion (139) from identity (136) and the analysis of elliptic orbits as specified in section (8), the trajectory of the orbital particle is found from (236).

4 Comparing quaternion and spinor notation for solving the Kepler problem

We have now arrived at the final discussion of this paper. In this section we will compare the two different methods of solving the Kepler problem, by either using quaternion notation, as demonstrated by J. Walvogel (2007) [13] or spinor notation as used by T. Bartsch (2003) [3]. We have seen that both methods arrive elegantly at the final differential equation that describes the motion of four uncoupled harmonic oscillators with a common frequency. The two methods have turned out to be very similar, which we will explain by showing that the underlying algebras of both methods are isomorphic to each other [4] [5]. After the proof of quaternions and spinors being isomorphic to each other, we will compare the similarities between both methods of regularization. Firstly, we will jump to the end of both methods and compare the final differential equations both methods arrive at. Thereafter, we will regard the fictitious time τ transformation as made in both methods of solving the Kepler problem. Thirdly, we will discuss both methods' need of a fourth dimension to describe a three dimensional problem; We will consider the additional degree of freedom and how it results from both methods and how this degree of freedom is used to constrain the problem. Lastly, we will compare how both arrive at a Kustaanheimo-Stiefel resembling transformation. We conclude this section by evaluating both methods and assessing which method is most suited in which situation.

4.1 Proof of quaternions and spinors being isomorphic

Both the four-dimensional quaternion and the four-dimensional spinor are often used for elegantly representing rotations in $\Re^3$. We will begin our comparison of the algebraic nature of quaternions and spinors by looking at their bases. A quaternion consists of a real part and a quaternion part; any quaternion $\mathbf{u} \in \mathbb{U}$ can be expressed in terms of the real unit 1 and the imaginary units i, j and k. Thus, the basis of a quaternion $\mathbf{u} \in \mathbb{U}$ is given by $(1, i, j, k)$. Similarly, a spinor consists of a (real) scalar part and a bivector part; any spinor $\mathbf{S} \in M_E$ can be expressed in terms of the (real) scalar unit 1 and the unit bivectors $\vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_3\vec{\sigma}_1$ and $\vec{\sigma}_2\vec{\sigma}_3$. Thus, the basis for a spinor $\mathbf{S} \in M_E$ is given by $(1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_3\vec{\sigma}_1, \vec{\sigma}_2\vec{\sigma}_3)$. So, both quaternions and spinors contain the same unit basis vector 1. We will now compare the quaternion part of the basis $(1, i, j, k)$ of quaternions to the bivector part of the basis $(1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_3\vec{\sigma}_1, \vec{\sigma}_2\vec{\sigma}_3)$ of spinors. We started our discourse of quaternion algebra with the governing relationship between the imaginary units i, j and k of quaternions (equation (64)):

$$i^2 = j^2 = k^2 = ijk = -1. \tag{237}$$

For a basis bivector $\vec{\sigma}_i\vec{\sigma}_j$ of a spinor, we can deduce the exact same relation, using (160):

$$(\vec{\sigma}_i\vec{\sigma}_j)^2 = \vec{\sigma}_i\vec{\sigma}_j\vec{\sigma}_i\vec{\sigma}_j = -\vec{\sigma}_i\vec{\sigma}_j\vec{\sigma}_j\vec{\sigma}_i = -\vec{\sigma}_i\vec{\sigma}_i = -1 \tag{238}$$

and

$$(\vec{\sigma}_1\vec{\sigma}_2)(\vec{\sigma}_3\vec{\sigma}_1)(\vec{\sigma}_2\vec{\sigma}_3) = \vec{\sigma}_2\vec{\sigma}_1\vec{\sigma}_1\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_3 = \vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_3 = -\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_2\vec{\sigma}_3 = -\vec{\sigma}_3\vec{\sigma}_3 = -1. \tag{239}$$

From (64) we obtained the well know multiplication rules for quaternions: $ij = -ji = k$, $jk = -kj = i$ $ki = -ik = j$. We can derive the same rules for the spinors basis unit bivectors, from equation (239):

$$(\vec{\sigma}_1\vec{\sigma}_2)(\vec{\sigma}_3\vec{\sigma}_1) = -\vec{\sigma}_1\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_1 = -\vec{\sigma}_3\vec{\sigma}_1\vec{\sigma}_1\vec{\sigma}_2 = -(\vec{\sigma}_3\vec{\sigma}_1)(\vec{\sigma}_1\vec{\sigma}_2) = -\vec{\sigma}_3\vec{\sigma}_2, \tag{240}$$

$$(\vec{\sigma}_2\vec{\sigma}_3)(\vec{\sigma}_1\vec{\sigma}_2) = -\vec{\sigma}_2\vec{\sigma}_1\vec{\sigma}_3\vec{\sigma}_2 = -\vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_2\vec{\sigma}_3 = -(\vec{\sigma}_1\vec{\sigma}_2)(\vec{\sigma}_2\vec{\sigma}_3) = -\vec{\sigma}_1\vec{\sigma}_3, \tag{241}$$

$$(\vec{\sigma}_3\vec{\sigma}_1)(\vec{\sigma}_2\vec{\sigma}_3) = -\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_1\vec{\sigma}_3 = -\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_3\vec{\sigma}_1 = -(\vec{\sigma}_2\vec{\sigma}_3)(\vec{\sigma}_3\vec{\sigma}_1) = -\vec{\sigma}_2\vec{\sigma}_1, \tag{242}$$

so the multiplication rules for the quaternion basis imaginary units i, j and k also hold for the basis $(1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_3\vec{\sigma}_1, \vec{\sigma}_2\vec{\sigma}_3)$ of spinors. Having illustrated this, we will now continue to prove that quaternions are isomorphic to spinors [4] [5].

We will prove first that the set of quaternions, together with the quaternion product, defines a group. Similarly, we will show the set of spinors (even multivectors), together with the geometric product, also forms a group. Finally, we will prove that these two groups are isomorphic.

In order to prove that these algebras form a group, we need to show they obey four properties:

Let G be the group and $\cdot$ the group operation.
1. *Closure*: for $a, b \in G : a \cdot b \in G$.
2. *Associativity*: for $a, b, c \in G : (ab)c = a(bc)$.
3. *Identity*: there is an $e \in G$, such that for any $a \in G : a \cdot e = e \cdot a = a$.
4. *Inverse*: for any $a \in G$, there exists an $a^{-1} \in G$, such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

We now present the proof that quaternions, together with the quaternion product form a group:
1. *Closure*: let $\mathbf{u}, \mathbf{v} \in \mathbb{U}$, then:

$$\mathbf{uv} = (u_0 + iu_1 + ju_2 + ku_3)(v_0 + iv_1 + jv_2 + kv_3) = (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3)... \tag{243}$$

$$... + i(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) + j(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) + k(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1),$$

which, as we can see, is again a quaternion, so $\mathbf{uv} \in \mathbb{U}$. This makes sense, because the units of quaternions form a closed group, any product of the units $1, i, j$ and k is equal again to one of the units $1, i, j$ and k, possibly with an added minus sign. Thus, any product of two quaternions can again be expressed in a sum of the units $1, i, j$ and k, multiplied individually by certain scalars.

2. *Associativity*: see proof 6.6 in *Appendix 1: Proofs*.

3. *Identity*: For any quaternion $\mathbf{u} \in \mathbb{U}$: $1\mathbf{u} = \mathbf{u}1 = \mathbf{u}$.

4. *Inverse*: Using equations (250) and (79), we have:

$$\mathbf{u^{-1}} = \frac{\bar{\mathbf{u}}}{\mathbf{u}\bar{\mathbf{u}}}, \tag{244}$$

$$\mathbf{u^{-1}u} = \frac{\bar{\mathbf{u}}\mathbf{u}}{\mathbf{u}\bar{\mathbf{u}}} = \frac{\mathbf{u}\bar{\mathbf{u}}}{\mathbf{u}\bar{\mathbf{u}}} = \mathbf{uu^{-1}} = 1, \tag{245}$$

so the set of quaternions $\mathbb{U}$, together with the quaternion product, form a group.

■

Proof that spinors, together with the geometric product form a group:
1. *Closure*: let $\mathbf{U}, \mathbf{V} \in M_E$, then:

$$\mathbf{UV} = (\langle\mathbf{U}\rangle_\mathbf{0} + \langle\mathbf{U}\rangle_\mathbf{2})(\langle\mathbf{V}\rangle_\mathbf{0} + \langle\mathbf{V}\rangle_\mathbf{2}) \tag{246}$$

$$= \langle\mathbf{V}\rangle_\mathbf{0}\langle\mathbf{U}\rangle_\mathbf{0} + \langle\mathbf{V}\rangle_\mathbf{0}\langle\mathbf{U}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{0}\langle\mathbf{V}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{2} \cdot \langle\mathbf{V}\rangle_\mathbf{2} + \langle\mathbf{U}\rangle_\mathbf{2} \times \langle\mathbf{V}\rangle_\mathbf{2}. \tag{247}$$

Using (179), we have:

$$\mathbf{UV} = (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3) + (u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2)\vec{\sigma}_1\vec{\sigma}_2...$$

$$... + (u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1)\vec{\sigma}_1\vec{\sigma}_3 + (u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1)\vec{\sigma}_2\vec{\sigma}_3, \tag{248}$$

which is a scalar-bivector sum, which is again a spinor, so $\mathbf{UV} \in M_E$. This makes sense, because the units of spinors form a closed group, any product of the units $1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_1\vec{\sigma}_3$ and $\vec{\sigma}_2\vec{\sigma}_3$ is equal again to one of the units $1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_1\vec{\sigma}_3$ and $\vec{\sigma}_2\vec{\sigma}_3$, possibly with an added minus sign. Thus, any product of two quaternions can again be expressed in a sum of the units $1, \vec{\sigma}_1\vec{\sigma}_2, \vec{\sigma}_1\vec{\sigma}_3$ and $\vec{\sigma}_2\vec{\sigma}_3$, multiplied individually by certain scalars.

2. *Associativity*: see proof 6.10 in *Appendix 1: Proofs*.
3. *Identity*: For any Spinor $\mathbf{U} \in \mathbb{U}$: $1\mathbf{U} = \mathbf{U}1 = \mathbf{U}$.
4. *Inverse*: Using equations (192) and (191), we have:

$$\mathbf{U^{-1}} = \frac{\mathbf{U}^\dagger}{|\mathbf{U}|^\mathbf{2}} = \frac{\mathbf{U}^\dagger}{\mathbf{UU}^\dagger}, \tag{249}$$

$$\mathbf{U^{-1}U} = \frac{\mathbf{U}^\dagger\mathbf{U}}{\mathbf{UU}^\dagger} = \frac{\mathbf{UU}^\dagger}{\mathbf{UU}^\dagger} = \mathbf{UU^{-1}} = 1, \tag{250}$$

so the set of spinors M_E, together with the geometric product, form a group.

■

We are now ready to show that the group of quaternions with the quaternion product is isomorfic to the group of spinors with the geometric product. Consider the function $f : \mathbb{U} \mapsto M_E$:

$$f(u_0 + iu_1 + ju_2 + ku_3) = u_0 + u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3, \tag{251}$$

with inverse:

$$f^{-1}(u_0 + u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3) = u_0. + iu_1 + ju_2 + ku_3 \tag{252}$$

To show that $f : \mathbb{U} \mapsto M_E$ is an isomorphism, we need to show that it is bijective and that $f : \mathbb{U} \mapsto M_E$ is a homomorphism.

Surjective: for every spinor $\mathbf{U} = a + b\vec{\sigma}_1\vec{\sigma}_2 + c\vec{\sigma}_1\vec{\sigma}_3 + d\vec{\sigma}_2\vec{\sigma}_3$, $a, b, c, d \in \Re$, there exists a quaternion, namely the quaternion $a + ib + jc + kd$, such that $f(a + ib + jc + kd) = a + b\vec{\sigma}_1\vec{\sigma}_2 + c\vec{\sigma}_1\vec{\sigma}_3 + d\vec{\sigma}_2\vec{\sigma}_3$

Injective: if $f(\mathbf{u}) = f(\mathbf{v}) = \mathbf{U} = a + b\vec{\sigma}_1\vec{\sigma}_2 + c\vec{\sigma}_1\vec{\sigma}_3 + d\vec{\sigma}_2\vec{\sigma}_3$, $a, b, c, d \in \Re$, then $\mathbf{u} = \mathbf{v} = f^{-1}(a + b\vec{\sigma}_1\vec{\sigma}_2 + c\vec{\sigma}_1\vec{\sigma}_3 + d\vec{\sigma}_2\vec{\sigma}_3) = a + ib + jc + kd$.

Thus, $f : \mathbb{U} \mapsto M_E$ is injective and surjective, so it is *bijective*.

■

Now, we will show that $f(\mathbf{uv}) = f(\mathbf{u})f(\mathbf{v})$, which completes the proof that quaternions and spinors are isomorphic:

Take $\mathbf{u}, \mathbf{v} \in \mathbb{U}$, $\mathbf{u} = u_0 + iu_1 + ju_2 + ku_3$ and $\mathbf{v} = v_0 + iv_1 + jv_2 + kv_3$, then:

$$\mathbf{uv} = (u_0 + iu_1 + ju_2 + ku_3)(v_0 + iv_1 + jv_2 + kv_3) = (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3)... \quad (253)$$

$$...+i(u_0v_1+u_1v_0+u_2v_3-u_3v_2)+j(u_0v_2+u_2v_0-u_1v_3+u_3v_1)+k(u_0v_3+u_3v_0+u_1v_2-u_2v_1),$$

so for $f(\mathbf{uv})$ we get:

$$f(\mathbf{uv}) = (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3) + (u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2)\vec{\sigma}_1\vec{\sigma}_2... \quad (254)$$

$$... + (u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1)\vec{\sigma}_1\vec{\sigma}_3 + (u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1)\vec{\sigma}_2\vec{\sigma}_3.$$

Now for $f(\mathbf{u})f(\mathbf{v})$:

$$f(\mathbf{u}) = u_0 + u_1\vec{\sigma}_1\vec{\sigma}_2 + u_2\vec{\sigma}_1\vec{\sigma}_3 + u_3\vec{\sigma}_2\vec{\sigma}_3 = \mathbf{U}, \quad (255)$$

$$f(\mathbf{v}) = v_0 + v_1\vec{\sigma}_1\vec{\sigma}_2 + v_2\vec{\sigma}_1\vec{\sigma}_3 + v_3\vec{\sigma}_2\vec{\sigma}_3 = \mathbf{V}, \quad (256)$$

which gives:

$$f(\mathbf{u})f(\mathbf{v}) = \mathbf{UV} = (u_0+u_1\vec{\sigma}_1\vec{\sigma}_2+u_2\vec{\sigma}_1\vec{\sigma}_3+u_3\vec{\sigma}_2\vec{\sigma}_3)(v_0+v_1\vec{\sigma}_1\vec{\sigma}_2+v_2\vec{\sigma}_1\vec{\sigma}_3+v_3\vec{\sigma}_2\vec{\sigma}_3) \quad (257)$$

$$= (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3) + (u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2)\vec{\sigma}_1\vec{\sigma}_2... \quad (258)$$

$$... + (u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1)\vec{\sigma}_1\vec{\sigma}_3 + (u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1)\vec{\sigma}_2\vec{\sigma}_3.$$

Now observe that equations (254) and (257) are identical, so $f(\mathbf{uv}) = f(\mathbf{u})f)\mathbf{v})$, which proves that the group of quaternions with the quaternion product is isomorphic to the group of spinors with the spinor product.

■

4.2 Similarities between both methods of regularization

We will now discuss an important characteristic of both methods regularizing the Kepler problem; the result they arrive upon. Both the quaternion and the spinor method end up with a differential equation that describes the motion of four uncoupled harmonic oscillators.

The quaternion method results in the differential equation (136):

$$2\mathbf{u}'' + h\mathbf{u} = 0, \tag{259}$$

which describes the motion of four uncoupled harmonic oscillators with common frequency $\omega := \sqrt{h/2}$.

The spinor method arrives upon the differential equation (236):

$$\mathbf{U}'' - 2H\mathbf{U} = 0, \tag{260}$$

which describes the motion of four uncoupled harmonic oscillators with common frequency $\omega := \sqrt{-2H}$ for $H < 0$.

By making the substitution $H = -\frac{1}{4}h$, we can see how the differential equation that resulted from the spinor method is equivalent to the differential equation we ended up with when using the quaternion method:

$$\mathbf{U}'' - 2H\mathbf{U} = \mathbf{U}'' - 2(-\frac{1}{4}h)\mathbf{U} = \mathbf{U}'' + \frac{1}{2}h\mathbf{U} = 2\mathbf{U}'' + h\mathbf{U} = 0 \tag{261}$$

and

$$\omega = \sqrt{-2H} = \sqrt{-2(-\frac{1}{4}h)} = \sqrt{h/2}. \tag{262}$$

Thus, we can see that both methods result in the same differential equation, describing the same harmonic oscillators with the same common frequency. This makes sense of course; the result should not depend on the choice of method of regularization.

Secondly, we regard the fictitious time τ transformation that is used in both methods. All methods of regularizing the Kepler problem in either two or three dimensions use some form of non-linear time scaling in order to deal with the singularity the Kepler problem has at its origin. However, it is interesting to see how both the quaternion method and the spinor method use almost the same fictitious time transformation $dt = rd\tau$ and $dt = 2rd\tau$, respectively. This resembles the fictitious time transformation as used by Levi-Civita (1920) [7]. This shows how essential the fictitious time transformation $dt = crd\tau$, $c \in \Re$ (c can be chosen arbitrarily as to simplify the equations) is to the regularization of the Kepler problem.

Thirdly, we will look at how both methods result in a degree of freedom (and thus need four dimensions to solve a three-dimensional problem) and how this degree of freedom is used to constrain the problem. In the quaternion method, the creation of the degree of freedom is straight forward; the mapping (95) maps the vector $\vec{x}$ into a product with vanishing k-component. In the spinor method the establishment of the degree of freedom also has a clear interpretation. The spinor method represents the vector $\vec{x}$ by a rotation dilatation of a fixed reference vector. However, rotations around the vector itself have no effect. Both methods use the degree of freedom impose a constraint on the problem in order to facilitate the differentiation of either the quaternion or spinor. For the quaternion method the constraint takes the form of a commutator relation. In the spinor method the constraint is used to eliminate trivector contributions, which means that rotations around the vector itself do not occur.

Lastly, we will now consider how both methods arrive upon a transformation that resembles the Kustaanheimo-Stiefel transformation [6].

The Kustaanheimo-Stiefel transformation maps $\Re^4$ to $\Re^3$ $(u_0, u_2, u_2, u_3) \mapsto (x_0, x_1, x_2)$ by:

$$\begin{aligned} x_0 &= u_0^2 + u_1^2 - u_2^2 - u_3^2 \\ x_1 &= 2(u_0 u_2 - u_1 u_3) \\ x_2 &= 2(u_0 u_3 + u_1 u_2). \end{aligned}$$

The quaternion transformation is given by:

$$\begin{aligned} x_0 &= u_0^2 - u_1^2 - u_2^2 + u_3^2 \\ x_1 &= 2u_0 u_1 - 2u_2 u_3 \\ x_2 &= 2u_0 u_2 + 2u_1 u_3. \end{aligned}$$

The spinor transformation is given by:

$$\begin{aligned} x &= u_1 u_3 - u_0 u_2 \\ y &= u_1 u_0 + u_2 u_3 \\ z &= \tfrac{1}{2}(u_0^2 - u_1^2 - u_2^2 + u_3^2). \end{aligned}$$

For the quaternion transformation we can see that the permutation of indices $1 \mapsto 2 \mapsto 3 \mapsto 1$ yields exactly the Kustaanheimo-Stiefel transformation. For the spinor transformation we need to multiply x, y and z by 2. We can see from how both the quaternion transformation and the spinor transformation resemble the Kustaanheimo-Stiefel transformation, how similar both transformations are to each other.

We conclude this section by evaluating both the quaternion method and the spinor method of solving the Kepler problem. Both methods solve the Kepler problem effectively. The method using quaternions uses a straight-forward, elegant transformation, as does the spinor method. Both methods should be readily understandable for physicists and mathematicians. Probably, for physicists using the spinor method is more accessible, as they will be more accustomed to using spinors than quaternions. For mathematicians on the other hand, I expect the reverse; they will probably be well-accustomed to using quaternions (depending on their field) and might find the quaternion method more elegant. As we have shown, quaternions and spinors turn out to be isomorphic, so there is no essential difference between both methods and hence the choice of method should depend solely on personal preference.

5 conclusion

To summarize, we have first introduced the Kepler problem; its formulation and applications. We have then shown how the Levi-Civita method regularizes the Kepler problem in two dimensions. Then, quaternion algebra was introduced and used to regularize the Kepler problem. Thereafter, geometric algebra was explained and used to regularize the Kepler problem as well. Finally, both methods and their results were compared and evaluated. We concluded that both methods are very well-suited for solving the Kepler problem. Based on my personal expectation, physicists will probably have more affinity for the spinor method, whereas mathematicians will probably prefer the quaternion method. The choice of method should be purely based on personal preference, as quaternions turn out to be isomorphic to spinors and thus both methods are essentially the same. As a suggestion for future research, it would be interesting to study how both methods compare when solving the pertubed Kepler problem, which arises for example when studying the orbit of an electron around a nucleus and applying an additional electric of magnetic field. Because of isomorphism, both methods will necessarily yield the same result, but there could be differences in practicallity. Personally, I find the quaternion method more elegant and practical.

References

[1] http://en.wikipedia.org/wiki/Newton's_laws_of_motion.

[2] http://en.wikipedia.org/wiki/Quaternion.

[3] T. Bartsch. The Kustaanheimo - Stiefel transformation in geometric algebra. *Journal of Physics A: Mathematical and General*, 36(25):6963, 2003.

[4] David Hestenes. Vectors, spinors, and complex numbers in classical and quantum physics. *Am. J. Phys*, 39(9):1013–27, 1971.

[5] John Kronsbein. Kinematics, quaternions, spinors and pauli's spin matrices. *American Journal of Physics*, 35:335, 1967.

[6] Paul E Kustaanheimo and E Stiefel. *Perturbation theory of Kepler motion based on spinor regularization.* Astronomical obervatory, 1965.

[7] Tullio Levi-Civita. Sur la régularisation du probleme des trois corps. *Acta mathematica*, 42(1):99–144, 1920.

[8] Pertti Lounesto. *Clifford algebras and spinors*, volume 286. Cambridge university press, 2001.

[9] W.A. Strauss. *Partial Differential Equations: An Introduction.* Wiley, 2008.

[10] K. F. Sundman. Recherches sur le probleme des trois corps. *Acta Societatis Scientificae Fennicae*, 34(6), 1907.

[11] Jörg Waldvogel. Order and chaos in satellite encounters. In *Chaotic Worlds: From Order to Disorder in Gravitational N-Body Dynamical Systems*, pages 231–251. Springer, 2006.

[12] Jörg Waldvogel. Quaternions and the perturbed Kepler problem. In *Periodic, Quasi-Periodic and Chaotic Motions in Celestial Mechanics: Theory and Applications*, pages 201–212. Springer, 2006.

[13] Jörg Waldvogel. Quaternions for regularizing celestial mechanics: the right way. *Celestial Mechanics and Dynamical Astronomy*, 102(1-3):149–162, 2008.

6 Appendix 1: Proofs

This section provides the proofs of various identities, as referred to in the paper.

6.1 Proof of the conservative vector field theorem

Statement (2) obviously implies (1), since it is a stronger version of one; for the force on the body to point away from or toward the origin, the force must be parallel or anti-parallel to the body's position vector $\vec{x}$, respectively and thus $F(\vec{x})$ must be equal to some scalar multiple of the position vector $\vec{x}$. This scalar multiple is given by $f(|\vec{x}|)$. Therefore, we only need to prove that (3) follows from (1) and (2) follows from (3). To show that (3) follows from (1), we need to show that $U(\vec{X}) = g(|\vec{X}|)$, because $F(\vec{x}) = -grad\ U(\vec{x})$ follows from the fact that the vector field $F(\vec{x})$ is conservative, according to relation 13 that was used to define conservative vector fields. Since g is some random function, all the relation $U(\vec{X}) = g(|\vec{x}|)$ tells us is that the potential U only depends on the distance $|\vec{x}|$ from the origin and thus it must be constant on a spehere, given by:

$$S_\alpha = \{\vec{X} \in \Re^n |\ \ |\vec{X}| = \alpha > 0\}. \tag{263}$$

We will show that U is constant on an curve γ on the sphere S_α, since any two points on the sphere S_α can be joined by a curve on the sphere S_α. For $J \subset \Re$ is an interval and $\gamma : J \mapsto S_\alpha$ is a smooth curve, then we will need to prove that the composition $U \circ \gamma$ is constant. Therefore, we will show its derivative is identically 0. By equation 7.4, the derivative equals:

$$\frac{d}{dt}U(\gamma(t)) = grad\ U(\gamma(t)) \cdot \dot{\gamma}(t). \tag{264}$$

Using equations 13 and 24, we have $grad\ U(\vec{X}) = -F(\vec{X}) = -\lambda(|\vec{X}|)\vec{X}$ becuase F is central according to (1), so:

$$\frac{d}{dt}U(\gamma(t)) = \lambda(\gamma(t))\gamma(t) \cdot \dot{\gamma}(t) \tag{265}$$

and using identity 20

$$= -\frac{\lambda(\gamma(t))}{2}\frac{d}{dt}|\gamma(t)|^2. \tag{266}$$

Now, $\frac{d}{dt}|\gamma(t)|^2 = 0$ because $|\gamma(t)| = constant = \alpha$, so:

$$\frac{d}{dt}U(\gamma(t)) = 0. \tag{267}$$

Thus, we have proven that (1) implies (3). We will now prove that (3) implies (2):

$$F(\vec{X}) = -grad\ U(\vec{X}) = -\left(\frac{\partial U}{\partial u_1}(\vec{X}), ..., \frac{\partial U}{\partial x_n}(\vec{X})\right) \tag{268}$$

and using the chain rule 7.4, we have:

$$= -\left(\frac{\partial g(|\vec{X}|)}{\partial |\vec{X}|}(\vec{X})\frac{\partial |\vec{X}|}{\partial x_1}, ..., \frac{\partial g(|\vec{X}|)}{\partial |\vec{X}|}(\vec{X})\frac{\partial |\vec{X}|}{\partial x_n}\right) \tag{269}$$

$$= -g'(|\vec{X}|)\left(\frac{2x_1}{2\sqrt{x_1^2+...+x_n^2}},...,\frac{2x_n}{2\sqrt{x_1^2+...+x_n^2}}\right), \tag{270}$$

$$F(\vec{X}) = -\frac{g'(|\vec{X}|)}{|\vec{X}|}\vec{X}. \tag{271}$$

Here, g' denotes the derivative of g with respect to $|\vec{X}|$. Thus, we have shown that (3) implies $F(\vec{X}) = f(|\vec{X}|)\vec{X}$, with $f(|\vec{X}|) = -g'(|\vec{X}|)/|\vec{X}|$. This completes the proof of theorem 2.1.

■

6.2 Nonlinearity of the Kepler force

For differential equation $\dot{\vec{v}} = -\frac{\vec{x}}{|\vec{x}|^3}$ (see 31) to be linear, we must be able to write it in the form $\mathbb{L}(\vec{x}) = 0$, where is a *linear operator*, which means the following must hold[9]:

$$\mathbb{L}(\vec{a}(t)+\vec{b}(t)) = \mathbb{L}(\vec{a}(t)) + \mathbb{L}(\vec{b}(t)) \quad \text{and} \quad \mathbb{L}(c\vec{a}(t)) = c\mathbb{L}(\vec{a}(t)). \tag{272}$$

Let $(\vec{a}(t)) = \dot{\vec{v}} + \frac{\vec{x}}{|\vec{x}|^3} = 0$, then:

$$\mathbb{L}(\vec{a}(t)+\vec{b}(t)) = \ddot{\vec{a}}(t) + \ddot{\vec{b}}(t) + \frac{\vec{a}(t)+\vec{b}(t)}{|\vec{a}(t)+\vec{b}(t)|^3} \tag{273}$$

$$\neq \ddot{\vec{a}}(t) + \ddot{\vec{b}}(t) + \frac{\vec{a}(t)}{|\vec{a}(t)|^3} + \frac{\vec{b}(t)}{|\vec{b}(t)|^3} = \mathbb{L}(\vec{a}(t)) + \mathbb{L}(\vec{b}(t)) \tag{274}$$

and

$$\mathbb{L}(c\vec{a}(t)) = c\ddot{\vec{a}} + \frac{c\vec{x}}{c^3|\vec{x}|^3} \tag{275}$$

$$\neq \ddot{\vec{a}} + \frac{\vec{x}}{|\vec{x}|^3} = c\mathbb{L}(\vec{a}(t)). \tag{276}$$

■

6.3 Linearity of the Kepler regularization

For differential equation $2\mathbf{u}'' + h\mathbf{u} = 0$ (see 61) to be linear, we must be able to write it in the form $\mathbb{L}(\mathbf{x}(\mathbf{t})) = 0$, where is a *linear operator*, which means the following must hold[9]:

$$\mathbb{L}(\mathbf{a}(t)+\mathbf{b}(t)) = \mathbb{L}(\mathbf{a}(t)) + \mathbb{L}(\mathbf{b}(t)) \quad \text{and} \quad \mathbb{L}(c\mathbf{a}(t)) = c\mathbb{L}(\mathbf{a}(t)). \tag{277}$$

Let $\mathbb{L}(\mathbf{a}(t)) = 2\mathbf{u}'' + h\mathbf{u} = 0$, then:

$$\mathbb{L}(\mathbf{a}(t)+\mathbf{b}(t)) = 2(\mathbf{a}(t)+\mathbf{b}(t))'' + h(\mathbf{a}(t)+\mathbf{b}(t)) \tag{278}$$

$$= 2(\mathbf{a}(t))'' + 2(\mathbf{b}(t))'' + h(\mathbf{a}(t)) + h(\mathbf{b}(t)) = \mathbb{L}(\mathbf{a}(t)) + \mathbb{L}(\mathbf{b}(t)), \tag{279}$$

$$\mathbb{L}(c\mathbf{a}(t)) = 2c\mathbf{a}(t)'' + ch\mathbf{a}(t) = c(2\mathbf{a}(\mathbf{t})'' + h\mathbf{a}(t)) = c\mathbb{L}(\mathbf{a}(t)). \tag{280}$$

■

6.4 Proof: Energy Integral of the Kepler force

Proof that

$$E(\vec{x},\vec{v}) = K(\vec{x}) + U(\vec{v}) = \frac{1}{2}|\vec{v}|^2 - \frac{1}{|\vec{x}|} \tag{281}$$

is the energy integral of

$$\dot{\vec{v}} = -\frac{\vec{x}}{|\vec{x}|^3} \text{ and } \dot{\vec{x}} = \vec{v}. \tag{282}$$

Since according to equation 23 $\dot{E} = 0$,

$$\frac{d}{dt}\left(\frac{1}{2}|\vec{v}|^2 - \frac{1}{|\vec{x}|}\right) = 0, \tag{283}$$

$$|\vec{v}| \cdot |\dot{\vec{v}}| + \frac{|\dot{\vec{x}}|}{|\vec{x}|^2} = 0, \tag{284}$$

$$|\vec{v}| \cdot \frac{d}{dt}\sqrt{v_1^2 + ... + v_n^2} + \frac{\frac{d}{dt}\sqrt{x_1^2 + ... + x_n^2}}{|\vec{x}|^2} = 0, \tag{285}$$

$$|\vec{v}| \cdot \frac{2v_1\dot{v}_1 + ... + 2v_n\dot{v}_n}{2|\vec{v}|} + \frac{\frac{2x_1\dot{x}_1 + ... + 2x_n\dot{x}_n}{2|\vec{x}|}}{|\vec{x}|^2} = 0, \tag{286}$$

$$\vec{v} \cdot \dot{\vec{v}} + \frac{\vec{x} \cdot \vec{v}}{|\vec{x}|^3} = 0, \tag{287}$$

$$\dot{\vec{v}} = -\frac{\vec{x}}{|\vec{x}|^3} = \vec{F}(\vec{x}). \tag{288}$$

■

6.5 Proof: distance to origin in elliptical orbit

$$r = |\mathbf{x}| = \sqrt{c^2 + 2ac\cos E + a^2\cos^2 E + b^2 \sin^2 E} \tag{289}$$

$$= \frac{1}{2}\sqrt{(A^4 - 2A^2B^2 + B^4) + 2(A^4 - B^4)\cos E + (A^4 + 2A^2B^2 + B^4)\cos^2 E + 4A^2B^2\sin^2 E} \tag{290}$$

$$= \sqrt{\frac{1}{4}(A^4 - 2A^2B^2 + B^4) + \frac{1}{2}(A^4 - B^4)\cos E + \frac{1}{4}(A^4 - 2A^2B^2 + B^4)\cos^2 E + A^2B^2}. \tag{291}$$

$$r = \sqrt{\frac{1}{4}(A^4 + 2A^2B^2 + B^4) + \frac{1}{2}(A^4 - B^4)\cos E + \frac{1}{4}(A^4 - 2A^2B^2 + B^4)\cos^2 E}, \tag{292}$$

$$r = \frac{A^2 + B^2}{2} + \frac{A^2 - B^2}{2} \cos E = a(1 + e \cos E), \tag{293}$$

$$r = \frac{A^2 + B^2}{2} + \frac{A^2 - B^2}{2} \cos E = a(1 + e \cos E). \tag{294}$$

6.6 Proof: associativity of the quaternion product

let $\mathbf{u}, \mathbf{v} \in \mathbb{U}$, then:

$$\mathbf{uv} = (u_0 + iu_1 + ju_2 + ku_3)(v_0 + iv_1 + jv_2 + kv_3) = (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3)... \tag{295}$$

$$...+i(u_0v_1+u_1v_0+u_2v_3-u_3v_2)+j(u_0v_2+u_2v_0-u_1v_3+u_3v_1)+k(u_0v_3+u_3v_0+u_1v_2-u_2v_1)$$

Take $\mathbf{uv} = \mathbf{x}$, so: $x_0 = u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3$, $x_1 = u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2$, $x_2 = u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1$ and $x_3 = u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1$. Then:

$$\mathbf{uv} = x_0 + ix_1 + jx_2 + kx_3, \tag{296}$$

$$\mathbf{xw} = (x_0+ix_1+jx_2+kx_3)(w_0+iw_1+jw_2+kw_3) = (x_0w_0-x_1w_1-x_2w_2-x_3w_3)... \tag{297}$$

$$...+i(x_0w_1+x_1w_0+x_2w_3-x_3w_2)+j(x_0w_2+x_2w_0-x_1w_3+x_3w_1)+k(x_0w_3+x_3w_0+x_1w_2-x_2w_1).$$

So the scalar multiples of the unit basis$(1, i, j, k)$ of the quaternion $(\mathbf{uv})\mathbf{w}$ are given *respectively* by:

For the real component, the scalar multiple of 1, we have:

$$x_0w_0-x_1w_1-x_2w_2-x_3w_3 = w_0(u_0v_0-u_1v_1-u_2v_2-u_3v_3)-w_1(u_0v_1+u_1v_0+u_2v_3-u_3v_2)...$$

$$... - w_2(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) - w_3(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \tag{298}$$

$$= w_0u_0v_0 - w_0u_1v_1 - w_0u_2v_2 - w_0u_3v_3 - w_1u_0v_1 - w_1u_1v_0 - w_1u_2v_3 + w_1u_3v_2...$$

$$... - w_2u_0v_2 - w_2u_2v_0 + w_2u_1v_3 - w_2u_3v_1 - w_3u_0v_3 - w_3u_3v_0 - w_3u_1v_2 + w_3u_2v_1. \tag{299}$$

For the i-component, the scalar multiple of i, we have:

$$x_0w_1+x_1w_0+x_2w_3-x_3w_2 = w_1(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_1+u_1v_0+u_2v_3-u_3v_2)...$$

$$... + w_3(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) - w_2(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \tag{300}$$

$$= w_1u_0v_0 - w_1u_1v_1 - w_1u_2v_2 - w_1u_3v_3 + w_0u_0v_1 + w_0u_1v_0 + w_0u_2v_3 - w_0u_3v_2...$$

$$... + w_3u_0v_2 + w_3u_2v_0 - w_3u_1v_3 + w_3u_3v_1 - w_2u_0v_3 - w_2u_3v_0 - w_2u_1v_2 + w_2u_2v_1. \tag{301}$$

For the j-component, the scalar multiple of j, we have:

$$x_0w_2+x_2w_0-x_1w_3+x_3w_1 = w_2(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_2+u_2v_0-u_1v_3+u_3v_1)...$$

$$... - w_3(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) + w_1(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \tag{302}$$

$$= w_2u_0v_0 - w_2u_1v_1 - w_2u_2v_2 - w_2u_3v_3 + w_0u_0v_2 + w_0u_2v_0 - w_0u_1v_3 + w_0u_3v_1...$$

$$... - w_3u_0v_1 - w_3u_1v_0 - w_3u_2v_3 + w_3u_3v_2 + w_1u_0v_3 + w_1u_3v_0 + w_1u_1v_2 - w_1u_2v_1. \tag{303}$$

For the k-component, the scalar multiple of k, we have

$$x_0w_3+x_3w_0+x_1w_2-x_2w_1 = w_3(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_3+u_3v_0+u_1v_2-u_2v_1)...$$

$$... + w_2(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) - w_1(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) \tag{304}$$

$$= w_3u_0v_0 - w_3u_1v_1 - w_3u_2v_2 - w_3u_3v_3 + w_0u_0v_3 + w_0u_3v_0 + w_0u_1v_2 - w_0u_2v_1...$$

$$... + w_2u_0v_1 + w_2u_1v_0 + w_2u_2v_3 - w_2u_3v_2 - w_1u_0v_2 - w_1u_2v_0 + w_1u_1v_3 - w_1u_3v_1. \tag{305}$$

Now for the right hand side of the equation $(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw})$:

$$\mathbf{vw} = (v_0+iv_1+jv_2+kv_3)(w_0+iw_1+jw_2+kw_3) = (v_0w_0-v_1w_1-v_2w_2-v_3w_3)... \tag{306}$$

$$...+i(v_0w_1+v_1w_0+v_2w_3-v_3w_2)+j(v_0w_2+v_2w_0-v_1w_3+v_3w_1)+k(v_0w_3+v_3w_0+v_1w_2-v_2w_1).$$

Take $\mathbf{vw} = \mathbf{y}$, so: $y_0 = v_0w_0-v_1w_1-v_2w_2-v_3w_3$, $y_1 = v_0w_1+v_1w_0+v_2w_3-v_3w_2$, $y_2 = v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1$ and $y_3 = v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1$. Then:

$$\mathbf{vw} = y_0 + iy_1 + jy_2 + ky_3, \tag{307}$$

$$\mathbf{uy} = (u_0 + iu_1 + ju_2 + ku_3)(y_0 + iy_1 + jy_2 + ky_3) = (y_0u_0 - y_1u_1 - y_2u_2 - y_3u_3)... \quad (308)$$

$$...+i(y_0u_1+y_1u_0+y_2u_3-y_3u_2)+j(y_0u_2+y_2u_0-y_1u_3+y_3u_1)+k(y_0u_3+y_3u_0+y_1u_2-y_2u_1).$$

Thus, the scalar multiples of the unit basis$(1, i, j, k)$ of the quaternion $\mathbf{u}(\mathbf{vw})$ are given *respectively* by:

For the real component, the scalar multiple of 1, we have:

$$y_0u_0-y_1u_1-y_2u_2-y_3u_3 = u_0(v_0w_0-v_1w_1-v_2w_2-v_3w_3)-u_1(v_0w_1+v_1w_0+v_2w_3-v_3w_2)...$$

$$... - u_2(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) - u_3(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (309)$$

$$= u_0v_0w_0 - u_0v_1w_1 - u_0v_2w_2 - u_0v_3w_3 - u_1v_0w_1 - u_1v_1w_0 - u_1v_2w_3 + u_1v_3w_2...$$

$$... - u_2v_0w_2 - u_2v_2w_0 + u_2v_1w_3 - u_2v_3w_1 - u_3v_0w_3 - u_3v_3w_0 - u_3v_1w_2 + u_3v_2w_1. \quad (310)$$

For the i-component, the scalar multiple of i, we have:

$$y_0u_1+y_1u_0+y_2u_3-y_3u_2 = u_1(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_1+v_1w_0+v_2w_3-v_3w_2)...$$

$$... + u_3(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) - u_2(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (311)$$

$$= u_1v_0w_0 - u_1v_1w_1 - u_1v_2w_2 - u_1v_3w_3 + u_0v_0w_1 + u_0v_1w_0 + u_0v_2w_3 - u_0v_3w_2...$$

$$... + u_3v_0w_2 + u_3v_2w_0 - u_3v_1w_3 + u_3v_3w_1 - u_2v_0w_3 - u_2v_3w_0 - u_2v_1w_2 + u_2v_2w_1. \quad (312)$$

For the j-component, the scalar multiple of j, we have:

$$y_0u_2+y_2u_0-y_1u_3+y_3u_1 = u_2(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_2+v_2w_0-v_1w_3+v_3w_1)...$$

$$... - u_3(v_0w_1 + v_1w_0 + v_2w_3 - v_3w_2) + u_1(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (313)$$

$$= u_2v_0w_0 - u_2v_1w_1 - u_2v_2w_2 - u_2v_3w_3 + u_0v_0w_2 + u_0v_2w_0 - u_0v_1w_3 + u_0v_3w_1...$$

$$... - u_3v_0w_1 - u_3v_1w_0 - u_3v_2w_3 + u_3v_3w_2 + u_1v_0w_3 + u_1v_3w_0 + u_1v_1w_2 - u_1v_2w_1. \quad (314)$$

For the k-component, the scalar multiple of k, we have:

$$y_0u_3+y_3u_0+y_1u_2-y_2u_1 = u_3(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_3+v_3w_0+v_1w_2-v_2w_1)...$$

$$... + u_2(v_0w_1 + v_1w_0 + v_2w_3 - v_3w_2) - u_1(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) \quad (315)$$

$$= u_3v_0w_0 - u_3v_1w_1 - u_3v_2w_2 - u_3v_3w_3 + u_0v_0w_3 + u_0v_3w_0 + u_0v_1w_2 - u_0v_2w_1...$$

$$... + u_2v_0w_1 + u_2v_1w_0 + u_2v_2w_3 - u_2v_3w_2 - u_1v_0w_2 - u_1v_2w_0 + u_1v_1w_3 - u_1v_3w_1. \quad (316)$$

Upon careful examination, we can see that equation 299 equals equation 310, equation 301 equals equation 312, equation 303 equals equation 314 and equation 305 equals equation 316 (the order of the scalars in the scalar products is different, as is the order of the terms in general, but, naturally, the scalar sum and scalar product are associative so the order does not matter). Thus we have proven associativity for the quaternion product: $(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw})$

■

.

6.7 The conjugation of a quaternion product

$$\overline{\mathbf{uv}} = \overline{(u_0 + iu_1 + ju_2 + ku_3)(v_0 + iv_1 + jv_2 + kv_3)} \quad (317)$$

$$= \overline{u_0v_0 + iu_0v_1 + ju_0v_2 + ku_0v_3 + iu_1v_0 - u_1v_1 + ku_1v_2 - ju_1v_3...} \quad (318)$$

$$\overline{... + ju_2v_0 - ku_2v_1 - u_2v_2 + iu_2v_3 + ku_3v_0 + ju_3v_1 - iu_3v_2 - u_3v_3}$$

$$= \overline{u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3 + i(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2)...} \quad (319)$$

$$\overline{... + j(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) + k(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1)}$$

$$= u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3 + i(-u_0v_1 - u_1v_0 - u_2v_3 + u_3v_2)... \quad (320)$$

$$... + j(-u_0v_2 - u_2v_0 + u_1v_3 - u_3v_1) + k(-u_0v_3 - u_3v_0 - u_1v_2 + u_2v_1)$$

$$= v_0u_0 - v_1u_1 - v_2u_2 - v_3u_3 - iv_1u_0 - iv_0u_1 + kjv_3u_2 + jkv_2u_3... \quad (321)$$

$$... - jv_2u_0 - jv_0u_2 + kiv_3u_1 + ikv_1u_3 - kv_3u_0 - kv_0u_3 + jiv_2u_1 + ijv_1u_2$$

$$= (v_0 - iv_1 - jv_2 - kv_3)(u_0 - iu_1 - ju_2 - ku_3) = \bar{\mathbf{v}}\bar{\mathbf{u}}. \quad (322)$$

6.8 The star conjugation of the quaternion product

$$(\mathbf{uv})^* = ((u_0 + iu_1 + ju_2 + ku_3)(v_0 + iv_1 + jv_2 + kv_3))^* \quad (323)$$

$$= (u_0v_0 + iu_0v_1 + ju_0v_2 + ku_0v_3 + iu_1v_0 - u_1v_1 + ku_1v_2 - ju_1v_3 ... \quad (324)$$

$$... + ju_2v_0 - ku_2v_1 - u_2v_2 + iu_2v_3 + ku_3v_0 + ju_3v_1 - iu_3v_2 - u_3v_3)^*$$

$$= (u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3 + i(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) ... \quad (325)$$

$$... + j(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) + k(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1))^*$$

$$= u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3 + i(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) ... \quad (326)$$

$$... + j(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) - k(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1)$$

$$= v_0u_0 - v_1u_1 - v_2u_2 - v_3u_3 + iv_1u_0 + iv_0u_1 - kjv_3u_2 - jkv_2u_3) ... \quad (327)$$

$$... + jv_2u_0 + v_0u_2 - kiv_3u_1 - ikv_1u_3) - ku_0v_3 - kv_0u_3 + jiv_2u_1 + ijv_1u_2)$$

$$= (v_0 + iv_1 + jv_2 - kv_3)(u_0 + iu_1 + ju_2 + ku_3) = (\mathbf{v}^*\mathbf{u}^*). \quad (328)$$

■

6.9 The cross-product notation of a bivector

$$\mathbf{I}\vec{a} \times \vec{b} = \mathbf{I}((a_2b_3 - a_3b_2)\vec{\sigma}_1 + (a_3b_1 - a_1b_3)\vec{\sigma}_2 + (a_1b_2 - a_2b_1)\vec{\sigma}_3) \quad (329)$$

$$= (a_2b_3 - a_3b_2)\vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_1 + (a_3b_1 - a_1b_3)\vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_2 + (a_1b_2 - a_2b_1)\vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_3 \quad (330)$$

$$= (a_2b_3 - a_3b_2)\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_1\vec{\sigma}_1 - (a_3b_1 - a_1b_3)\vec{\sigma}_1\vec{\sigma}_3\vec{\sigma}_2\vec{\sigma}_2 + (a_1b_2 - a_2b_1)\vec{\sigma}_1\vec{\sigma}_2\vec{\sigma}_3\vec{\sigma}_3 \quad (331)$$

$$= (a_2b_3 - a_3b_2)\vec{\sigma}_2\vec{\sigma}_3 + (a_3b_1 - a_1b_3)\vec{\sigma}_3\vec{\sigma}_1 + (a_1b_2 - a_2b_1)\vec{\sigma}_1\vec{\sigma}_2 = \vec{a} \wedge \vec{b}. \quad (332)$$

■

6.10 Proof: associativity of the geometric product of spinors (even multivectors)

Proof: let $\mathbf{U}, \mathbf{V} \in \mathbb{U}$, then:

$$\mathbf{UV} = (u_0+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}u_1+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}u_2+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}u_3)(v_0+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}v_1+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}v_2+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}v_3) = (u_0v_0-u_1v_1-u_2v_2-u_3v_3)... \quad (333)$$

$$...+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}(u_0v_1+u_1v_0+u_2v_3-u_3v_2)+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}(u_0v_2+u_2v_0-u_1v_3+u_3v_1)+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}(u_0v_3+u_3v_0+u_1v_2-u_2v_1).$$

Take $\mathbf{UV} = \mathbf{X}$, so: $x_0 = u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3$, $x_1 = u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2$, $x_2 = u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1$ and $x_3 = u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1$. Then:

$$\mathbf{UV} = x_0 + \tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}x_1 + \tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}x_2 + \tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}x_3 \quad (334)$$

$$\mathbf{XW} = (x_0+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}x_1+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}x_2+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}x_3)(w_0+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}w_1+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}w_2+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}w_3) = (x_0w_0-x_1w_1-x_2w_2-x_3w_3)... \quad (335)$$

$$...+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}(x_0w_1+x_1w_0+x_2w_3-x_3w_2)+\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}(x_0w_2+x_2w_0-x_1w_3+x_3w_1)+\tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3}(x_0w_3+x_3w_0+x_1w_2-x_2w_1).$$

So the scalar multiples of the unit basis$(1, \tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}, \tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{3}, \tilde{\sigma}_\mathbf{2}\tilde{\sigma}_\mathbf{3})$ of the spinor $(\mathbf{UV})\mathbf{W}$ are given *respectively* by:

For the real component, the scalar multiple of 1, we have:

$$x_0w_0-x_1w_1-x_2w_2-x_3w_3 = w_0(u_0v_0-u_1v_1-u_2v_2-u_3v_3)-w_1(u_0v_1+u_1v_0+u_2v_3-u_3v_2)...$$

$$... - w_2(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) - w_3(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \quad (336)$$

$$= w_0u_0v_0 - w_0u_1v_1 - w_0u_2v_2 - w_0u_3v_3 - w_1u_0v_1 - w_1u_1v_0 - w_1u_2v_3 + w_1u_3v_2...$$

$$... - w_2u_0v_2 - w_2u_2v_0 + w_2u_1v_3 - w_2u_3v_1 - w_3u_0v_3 - w_3u_3v_0 - w_3u_1v_2 + w_3u_2v_1. \quad (337)$$

For the $\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}$-component, the scalar multiple of $\tilde{\sigma}_\mathbf{1}\tilde{\sigma}_\mathbf{2}$, we have:

$$x_0w_1+x_1w_0+x_2w_3-x_3w_2 = w_1(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_1+u_1v_0+u_2v_3-u_3v_2)...$$

$$... + w_3(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) - w_2(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \quad (338)$$

$$= w_1u_0v_0 - w_1u_1v_1 - w_1u_2v_2 - w_1u_3v_3 + w_0u_0v_1 + w_0u_1v_0 + w_0u_2v_3 - w_0u_3v_2...$$

$$... + w_3u_0v_2 + w_3u_2v_0 - w_3u_1v_3 + w_3u_3v_1 - w_2u_0v_3 - w_2u_3v_0 - w_2u_1v_2 + w_2u_2v_1. \quad (339)$$

For the $\tilde{\sigma}_1\tilde{\sigma}_3$-component, the scalar multiple of $\tilde{\sigma}_1\tilde{\sigma}_3$, we have:

$$x_0w_2+x_2w_0-x_1w_3+x_3w_1 = w_2(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_2+u_2v_0-u_1v_3+u_3v_1)...$$

$$... - w_3(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) + w_1(u_0v_3 + u_3v_0 + u_1v_2 - u_2v_1) \quad (340)$$

$$= w_2u_0v_0 - w_2u_1v_1 - w_2u_2v_2 - w_2u_3v_3 + w_0u_0v_2 + w_0u_2v_0 - w_0u_1v_3 + w_0u_3v_1...$$

$$... - w_3u_0v_1 - w_3u_1v_0 - w_3u_2v_3 + w_3u_3v_2 + w_1u_0v_3 + w_1u_3v_0 + w_1u_1v_2 - w_1u_2v_1. \quad (341)$$

For the $\tilde{\sigma}_2\tilde{\sigma}_3$-component, the scalar multiple of $\tilde{\sigma}_2\tilde{\sigma}_3$, we have:

$$x_0w_3+x_3w_0+x_1w_2-x_2w_1 = w_3(u_0v_0-u_1v_1-u_2v_2-u_3v_3)+w_0(u_0v_3+u_3v_0+u_1v_2-u_2v_1)...$$

$$... + w_2(u_0v_1 + u_1v_0 + u_2v_3 - u_3v_2) - w_1(u_0v_2 + u_2v_0 - u_1v_3 + u_3v_1) \quad (342)$$

$$= w_3u_0v_0 - w_3u_1v_1 - w_3u_2v_2 - w_3u_3v_3 + w_0u_0v_3 + w_0u_3v_0 + w_0u_1v_2 - w_0u_2v_1...$$

$$... + w_2u_0v_1 + w_2u_1v_0 + w_2u_2v_3 - w_2u_3v_2 - w_1u_0v_2 - w_1u_2v_0 + w_1u_1v_3 - w_1u_3v_1. \quad (343)$$

Now for the right hand side of the equation $(\mathbf{UV})\mathbf{W} = \mathbf{U}(\mathbf{VW})$:

$$\mathbf{VW} = (v_0+\tilde{\sigma}_1\tilde{\sigma}_2v_1+\tilde{\sigma}_1\tilde{\sigma}_3v_2+\tilde{\sigma}_2\tilde{\sigma}_3v_3)(w_0+\tilde{\sigma}_1\tilde{\sigma}_2w_1+\tilde{\sigma}_1\tilde{\sigma}_3w_2+\tilde{\sigma}_2\tilde{\sigma}_3w_3) = (v_0w_0-v_1w_1-v_2w_2-v_3w_3)... \quad (344)$$

$$...+\tilde{\sigma}_1\tilde{\sigma}_2(v_0w_1+v_1w_0+v_2w_3-v_3w_2)+\tilde{\sigma}_1\tilde{\sigma}_3(v_0w_2+v_2w_0-v_1w_3+v_3w_1)+\tilde{\sigma}_2\tilde{\sigma}_3(v_0w_3+v_3w_0+v_1w_2-v_2w_1).$$

Take $\mathbf{VW} = \mathbf{Y}$, so: $y_0 = v_0w_0-v_1w_1-v_2w_2-v_3w_3$, $y_1 = v_0w_1+v_1w_0+v_2w_3-v_3w_2$, $y_2 = v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1$ and $y_3 = v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1$. Then:

$$\mathbf{VW} = y_0 + \tilde{\sigma}_1\tilde{\sigma}_2y_1 + \tilde{\sigma}_1\tilde{\sigma}_3y_2 + \tilde{\sigma}_2\tilde{\sigma}_3y_3, \quad (345)$$

$$\mathbf{UY} = (u_0+\tilde{\sigma}_1\tilde{\sigma}_2u_1+\tilde{\sigma}_1\tilde{\sigma}_3u_2+\tilde{\sigma}_2\tilde{\sigma}_3u_3)(y_0+\tilde{\sigma}_1\tilde{\sigma}_2y_1+\tilde{\sigma}_1\tilde{\sigma}_3y_2+\tilde{\sigma}_2\tilde{\sigma}_3y_3) = (y_0u_0-y_1u_1-y_2u_2-y_3u_3)... \quad (346)$$

$$...+\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{2}}(y_0u_1+y_1u_0+y_2u_3-y_3u_2)+\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{3}}(y_0u_2+y_2u_0-y_1u_3+y_3u_1)+\tilde{\sigma}_{\mathbf{2}}\tilde{\sigma}_{\mathbf{3}}(y_0u_3+y_3u_0+y_1u_2-y_2u_1).$$

So the scalar multiples of the unit basis$(1, \tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{2}}, \tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{3}}, \tilde{\sigma}_{\mathbf{2}}\tilde{\sigma}_{\mathbf{3}})$ of the spinor $\mathbf{U}(\mathbf{VW})$ are given *respectively* by:

For the scalar component, the scalar multiple of 1, we have:

$$y_0u_0-y_1u_1-y_2u_2-y_3u_3 = u_0(v_0w_0-v_1w_1-v_2w_2-v_3w_3)-u_1(v_0w_1+v_1w_0+v_2w_3-v_3w_2)...$$

$$... - u_2(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) - u_3(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (347)$$

$$= u_0v_0w_0 - u_0v_1w_1 - u_0v_2w_2 - u_0v_3w_3 - u_1v_0w_1 - u_1v_1w_0 - u_1v_2w_3 + u_1v_3w_2...$$

$$... - u_2v_0w_2 - u_2v_2w_0 + u_2v_1w_3 - u_2v_3w_1 - u_3v_0w_3 - u_3v_3w_0 - u_3v_1w_2 + u_3v_2w_1. \quad (348)$$

For the $\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{2}}$-component, the scalar multiple of $\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{2}}$, we have:

$$y_0u_1+y_1u_0+y_2u_3-y_3u_2 = u_1(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_1+v_1w_0+v_2w_3-v_3w_2)...$$

$$... + u_3(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) - u_2(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (349)$$

$$= u_1v_0w_0 - u_1v_1w_1 - u_1v_2w_2 - u_1v_3w_3 + u_0v_0w_1 + u_0v_1w_0 + u_0v_2w_3 - u_0v_3w_2...$$

$$... + u_3v_0w_2 + u_3v_2w_0 - u_3v_1w_3 + u_3v_3w_1 - u_2v_0w_3 - u_2v_3w_0 - u_2v_1w_2 + u_2v_2w_1. \quad (350)$$

For the $\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{3}}$-component, the scalar multiple of $\tilde{\sigma}_{\mathbf{1}}\tilde{\sigma}_{\mathbf{3}}$, we have:

$$y_0u_2+y_2u_0-y_1u_3+y_3u_1 = u_2(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_2+v_2w_0-v_1w_3+v_3w_1)...$$

$$... - u_3(v_0w_1 + v_1w_0 + v_2w_3 - v_3w_2) + u_1(v_0w_3 + v_3w_0 + v_1w_2 - v_2w_1) \quad (351)$$

$$= u_2v_0w_0 - u_2v_1w_1 - u_2v_2w_2 - u_2v_3w_3 + u_0v_0w_2 + u_0v_2w_0 - u_0v_1w_3 + u_0v_3w_1...$$

$$... - u_3v_0w_1 - u_3v_1w_0 - u_3v_2w_3 + u_3v_3w_2 + u_1v_0w_3 + u_1v_3w_0 + u_1v_1w_2 - u_1v_2w_1. \quad (352)$$

For the $\tilde{\sigma}_{\mathbf{2}}\tilde{\sigma}_{\mathbf{3}}$-component, the scalar multiple of $\tilde{\sigma}_{\mathbf{2}}\tilde{\sigma}_{\mathbf{3}}$, we have:

$$y_0u_3+y_3u_0+y_1u_2-y_2u_1 = u_3(v_0w_0-v_1w_1-v_2w_2-v_3w_3)+u_0(v_0w_3+v_3w_0+v_1w_2-v_2w_1)...$$

$$... + u_2(v_0w_1 + v_1w_0 + v_2w_3 - v_3w_2) - u_1(v_0w_2 + v_2w_0 - v_1w_3 + v_3w_1) \quad (353)$$

$$= u_3v_0w_0 - u_3v_1w_1 - u_3v_2w_2 - u_3v_3w_3 + u_0v_0w_3 + u_0v_3w_0 + u_0v_1w_2 - u_0v_2w_1...$$

$$... + u_2v_0w_1 + u_2v_1w_0 + u_2v_2w_3 - u_2v_3w_2 - u_1v_0w_2 - u_1v_2w_0 + u_1v_1w_3 - u_1v_3w_1. \quad (354)$$

Upon careful examination, we can see that equation 337 equals equation 348, equation 339 equals equation 350, equation 341 equals equation 352 and equation 343 equals equation 354 (the order of the scalars in the scalar products is different, as is the order of the terms in general, but, naturally, the scalar sum and scalar product are associative so the order does not matter). Thus we have proven associativity for the spinor product: $(\mathbf{UV})\mathbf{W} = \mathbf{U}(\mathbf{VW})$.

■

6.11 The reversion of the geometric product of spinors

Let $\mathbf{A}, \mathbf{B} \in M_E$, $\mathbf{A} = a_0+a_1\vec{\sigma}_1\vec{\sigma}_2+a_2\vec{\sigma}_1\vec{\sigma}_3+a_3\vec{\sigma}_2\vec{\sigma}_3$ and $\mathbf{B} = a_0+b_1\vec{\sigma}_1\vec{\sigma}_2+b_2\vec{\sigma}_1\vec{\sigma}_3+b_3\vec{\sigma}_2\vec{\sigma}_3$, then:

$$(\mathbf{AB})^\dagger = (((\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{A}\rangle_\mathbf{2})(\langle\mathbf{B}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle_\mathbf{2}))^\dagger \quad (355)$$

$$= (\langle\mathbf{A}\rangle_\mathbf{0}\langle\mathbf{B}\rangle_\mathbf{0} + \langle\mathbf{A}\rangle_\mathbf{0}\langle\mathbf{B}\rangle_\mathbf{2} + \langle\mathbf{A}\rangle_\mathbf{2}\langle\mathbf{B}\rangle_\mathbf{0} + \langle\mathbf{A}\rangle_\mathbf{2}\langle\mathbf{B}\rangle_\mathbf{2})^\dagger \quad (356)$$

$$= \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle^\dagger_\mathbf{2}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle^\dagger_\mathbf{2}...$$

$$... + ((\langle\mathbf{A}\rangle_\mathbf{2} \cdot \langle\mathbf{B}\rangle_\mathbf{2} + \langle\mathbf{A}\rangle_\mathbf{2} \times \langle\mathbf{B}\rangle_\mathbf{2})^\dagger \quad (357)$$

$$= \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle^\dagger_\mathbf{2}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle^\dagger_\mathbf{2}...$$

$$... + ((\langle\mathbf{A}\rangle_\mathbf{2} \cdot \langle\mathbf{B}\rangle_\mathbf{2})^\dagger + ((\langle\mathbf{A}\rangle_\mathbf{2} \times \langle\mathbf{B}\rangle_\mathbf{2})^\dagger. \quad (358)$$

$\langle\mathbf{B}\rangle_\mathbf{2} \cdot \langle\mathbf{A}\rangle_\mathbf{2}$ is a scalar, so is does not change under revision (and I also use that it is commutative):

$$= \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle^\dagger_\mathbf{2}\langle\mathbf{A}\rangle_\mathbf{0} + \langle\mathbf{B}\rangle_\mathbf{0}\langle\mathbf{A}\rangle^\dagger_\mathbf{2}...$$

$$... + \langle\mathbf{B}\rangle_\mathbf{2} \cdot \langle\mathbf{A}\rangle_\mathbf{2} + ((a_2b_3 - a_3b_2)\vec{\sigma}_1\vec{\sigma}_2 - (a_1b_3 - a_3b_1)\vec{\sigma}_1\vec{\sigma}_3 + (a_1b_2 - a_2b_1)\vec{\sigma}_2\vec{\sigma}_3)^\dagger \quad (359)$$

$$= \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{2}}^{\dagger} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{2}}^{\dagger} ...$$

$$... + \langle \mathbf{B} \rangle_{\mathbf{2}} \cdot \langle \mathbf{A} \rangle_{\mathbf{2}} + (a_2 b_3 - a_3 b_2) \vec{\sigma}_2 \vec{\sigma}_1 - (a_1 b_3 - a_3 b_1) \vec{\sigma}_3 \vec{\sigma}_1 + (a_1 b_2 - a_2 b_1) \vec{\sigma}_3 \vec{\sigma}_2 \tag{360}$$

$$= \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{2}}^{\dagger} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{2}}^{\dagger} ...$$

$$... + \langle \mathbf{B} \rangle_{\mathbf{2}} \cdot \langle \mathbf{A} \rangle_{\mathbf{2}} + (b_2 a_3 - b_3 a_2) \vec{\sigma}_1 \vec{\sigma}_2 - (b_1 a_3 - b_3 a_1) \vec{\sigma}_1 \vec{\sigma}_3 + (b_1 a_2 - b_2 a_1) \vec{\sigma}_2 \vec{\sigma}_3 \tag{361}$$

$$= \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{2}}^{\dagger} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{2}}^{\dagger} ...$$

$$... + \langle \mathbf{B} \rangle_{\mathbf{2}} \cdot \langle \mathbf{A} \rangle_{\mathbf{2}} + (-b_1 \vec{\sigma}_1 \vec{\sigma}_2 - b_2 \vec{\sigma}_1 \vec{\sigma}_3 + b_3 \vec{\sigma}_2 \vec{\sigma}_3) \times (-a_1 \vec{\sigma}_1 \vec{\sigma}_2 - a_2 \vec{\sigma}_1 \vec{\sigma}_3 - a_3 \vec{\sigma}_2 \vec{\sigma}_3) \tag{362}$$

$$= (\langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{0}})^{\dagger} + \langle \mathbf{B} \rangle_{\mathbf{2}}^{\dagger} \langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{0}} \langle \mathbf{A} \rangle_{\mathbf{2}}^{\dagger} + (\langle \mathbf{B} \rangle_{\mathbf{2}} \cdot \langle \mathbf{A} \rangle_{\mathbf{2}})^{\dagger} + \langle \mathbf{B} \rangle_{\mathbf{2}}^{\dagger} \times \langle \mathbf{A} \rangle_{\mathbf{2}}^{\dagger} \tag{363}$$

$$= (\langle \mathbf{B} \rangle_{\mathbf{0}} + \langle \mathbf{B} \rangle_{\mathbf{2}})^{\dagger} (\langle \mathbf{A} \rangle_{\mathbf{0}} + \langle \mathbf{A} \rangle_{\mathbf{2}})^{\dagger} = \mathbf{B}^{\dagger} \mathbf{A}^{\dagger}. \tag{364}$$

■

6.12 The modulus of a spinor

Let $\mathbf{S} \in M_E$, $\mathbf{S} = s_0 + s_1 \vec{\sigma}_1 \vec{\sigma}_2 + s_2 \vec{\sigma}_1 \vec{\sigma}_3 + s_3 \vec{\sigma}_2 \vec{\sigma}_3$, then:

$$|\mathbf{S}|^{\mathbf{2}} = \mathbf{S}\mathbf{S}^{\dagger} = (\langle \mathbf{S} \rangle_{\mathbf{0}} + \langle \mathbf{S} \rangle_{\mathbf{2}})(\langle \mathbf{S} \rangle_{\mathbf{0}} + \langle \mathbf{S} \rangle_{\mathbf{2}})^{\dagger} = (\langle \mathbf{S} \rangle_{\mathbf{0}} + \langle \mathbf{S} \rangle_{\mathbf{2}})(\langle \mathbf{S} \rangle_{\mathbf{0}} - \langle \mathbf{S} \rangle_{\mathbf{2}}) \tag{365}$$

$$= \langle \mathbf{S} \rangle_{\mathbf{0}} \langle \mathbf{S} \rangle_{\mathbf{0}} - \langle \mathbf{S} \rangle_{\mathbf{0}} \langle \mathbf{S} \rangle_{\mathbf{2}} + \langle \mathbf{S} \rangle_{\langle} \mathbf{S} \rangle_{\mathbf{0}} + \langle \mathbf{S} \rangle_{\mathbf{2}} \langle \mathbf{S} \rangle_{\mathbf{2}} \tag{366}$$

$$= \langle \mathbf{S} \rangle_{\mathbf{0}}^{\mathbf{2}} + \langle \mathbf{S} \rangle_{\mathbf{2}} \cdot \langle \mathbf{S} \rangle_{\mathbf{2}} + \langle \mathbf{S} \rangle_{\mathbf{2}} \times \langle \mathbf{S} \rangle_{\mathbf{2}} = \langle \mathbf{S} \rangle_{\mathbf{0}}^{\mathbf{2}} + \langle \mathbf{S} \rangle_{\mathbf{2}} \cdot \langle \mathbf{S} \rangle_{\mathbf{2}} \tag{367}$$

$$= s_0^2 - s_1^2 - s_2^2 - s_3^2. \tag{368}$$

■

7 Appendix 2: Multivariable Calculus

We assume knowlegde of some basic concepts from multivariable calculus in this thesis, which, for completeness, are outlined here:

Definition 7.1. The *dot product* or *scalar product* of two vectors $\vec{x}, \vec{y} \in \Re^n$, denoted as $\vec{x} \cdot \vec{y}$ is defined as:

$$\vec{x} \cdot \vec{y} = \Sigma_{i=1}^n x_i y_i, \tag{369}$$

with $\vec{x} = (x_1, ..., x_n)$ and $\vec{y} = (y_1, ..., y_n)$. Note that:

$$\vec{x} \cdot \vec{x} = \Sigma_{i=1}^n x_i^2 = |\vec{x}|^2. \tag{370}$$

Definition 7.2. The *gradient* of a function $g : \Re^n \mapsto \Re$, denoted as *grad* g, is defined by:

$$grad \;\; g(\vec{x}) = \left(\frac{\partial g}{\partial x_i}(\vec{x}), ..., \frac{\partial g}{\partial x_n}(\vec{x}) \right), \tag{371}$$

which gives a vector field on $\Re^n$.

Definition 7.3. The *inner* product rule for two smooth functions f, g: $I \mapsto \Re^n$ is given by:

$$\frac{d}{dt}(f \cdot g) = \dot{f} \cdot g + f \cdot \dot{g}, \tag{372}$$

where $(\dot{}) = d/dt$.

Definition 7.4. The chain rule, applied to a composition $g \circ F$ of two functions $F : \Re \mapsto \Re^n$ and $g : \Re^n \mapsto \Re$, yields:

$$\frac{d}{dt} g(F(t)) = grad \;\; g(F(t)) \cdot \dot{F}(t) \tag{373}$$

$$= \Sigma_{i=1}^n \frac{\partial g}{\partial F_i}(F(t)) \frac{dF_i}{dt}. \tag{374}$$

We will also consider the *cross product*:

Definition 7.5. The cross product of two vectors $\vec{u}, \vec{v} \in \Re^3$, denoted as $\vec{u} \times \vec{v}$, is given by:

$$\vec{u} \times \vec{v} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 u_2 - u_2 u_1) \tag{375}$$

$$= -\vec{v} \times \vec{u} = |\vec{u}||\vec{v}| \sin \theta \hat{N} \in \Re^3. \tag{376}$$

Here, θ denotes the angle between $\vec{u}$ and $\vec{v}$ and $\hat{N}$ denotes the unit vector in the positive direction (according to the right hand rule) of $\vec{u} \times \vec{v}$ ($\vec{N}$ is somtimes called the *normal*), so $|\vec{u}||\vec{v}| \sin \theta$ gives the magnitude of $\vec{u} \times \vec{v}$ ($\vec{u} \times \vec{v}$ points perpendicular to both $\vec{u}$ and $\vec{v}$ and is oriented according to the right hand rule).

From definition 7.5 follows that $\vec{u} \times \vec{v} = 0$ only when $\vec{u}$ and $\vec{v}$ are parallel. If $\vec{u} \times \vec{v} \neq 0$, then $\vec{u} \times \vec{v}$ is perpendicular tot the plane in which the vectors $\vec{u}$ and $\vec{v}$ lie.

Definition 7.6. The "cross" product rule is given by:

$$\frac{d}{dt}(\vec{u} \times \vec{v}) = \dot{\vec{u}} \times \vec{v} + \vec{u} \times \dot{\vec{v}}. \tag{377}$$

Lastly, we will frequently use the identity:

$$\frac{d}{dt}|\vec{X}|^2 = \frac{d}{dt}\sqrt{x_1^2 + ... + x_n^2}^2 = \left(\frac{d}{dx_1}x_1^2\frac{dx_1}{dt} + ... + \frac{d}{dx_n}x_n^2\frac{dx_n}{dt}\right) = 2\vec{X} \cdot \dot{\vec{X}}. \tag{378}$$

8 Appendix 3: The resulting orbital motion from the Kepler problem

Here we will discuss the resulting orbit from equation (63). Two cases need to be distinguished: Depending on whether or not the orbital body moves directly towards the central body, the orbital body will not or will, respectively, possess an angular momentum with respect to the central body. If the orbital body possesses no angular momentum with respect to the central body, the solution of the Kepler problem is very simple: The resulting trajectory of the orbital body is a straight path the central body, continuously accelerating until the two bodies collide. However, if the orbital body does have a very small amount of angular momentum, the resulting trajectory will just barely miss the central body (the two bodies are considered to be point masses), but will still closely resemble a collision trajectory. Thus, such a trajectory can be considered an approximation to a collision trajectory, but withouth the actual occurrence of the collision and the oribital body instead continuing its trajectory after reaching the origin. Aside from the very small amount of angular momentum, the resulting "orbit" is a one dimensional harmonic oscilator.

If the orbital body does possess (significant) angular momentum and no additional forces are present, the resulting motion is two-dimensional, in the plane spanned by the orbital body's velocity vector and the separation vector between the orbital body and the central body. The solution of the trajectory of the orbital body in the attractive field of the central particle is given by:

$$\mathbf{x} = \frac{A^2 - B^2}{2} + \frac{A^2 + B^2}{2}\cos 2\omega\tau + iAB\sin 2\omega\tau, \tag{379}$$

where $\mathbf{x}$ is the position vector, A and B are determined by the initial position and velocity vectors; the so called initial conditions, τ is non-linearly scaled time variable (dependent on r) and $\omega = \sqrt{h/2}$. The facor $E := 2\omega\tau$, called the *eccentric anomaly* is one of the most important parameters in defining Keplerian motions. Or more clearly, when substituting $a = \frac{A^2+B^2}{2}$, $b = AB$ and $c = \frac{A^2-B^2}{2}$:

$$\mathbf{x} = c + a\cos 2\omega\tau + ib\sin 2\omega\tau. \tag{380}$$

This describes an elliptic Keplerian orbit, as shown in figure 2 [12]. An ellipse is a closed curve that has two perpendicular symmetry axes; its major semi-axis and minor semi-axis. The (horizontal) major semi-axis is the longest possible line segment between two points on the elliptic orbit and is given by a (the amplitude of the cosine). The (vertical) minor semi-axis, which is perpendicular to the major semi-axis, is given by the

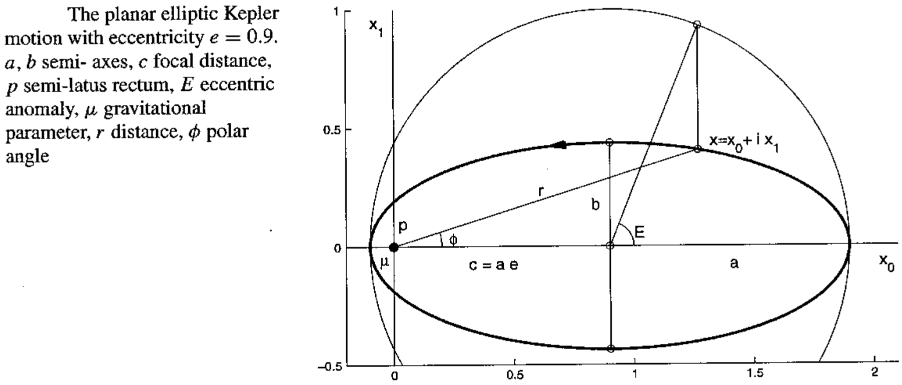

Figure 2: Elliptic Keplerian Orbit

shortest distance between two point on the elliptic orbit on opposite sides of the major semi-axis and is given by b (the amplitude of the sine). The center of the elliptic orbit lies on the intersection of the major and minor semi-axes. The distance of the central particle, the origin of our system, to the center of the elliptic orbit is given by c (the spacial shift of a system centered at the origin). An ellipse has two unique points, its foci, which lie on the major semi-axis at an equal distance from the center. The two foci are the only points for which the following relation holds: The sum of the distances of any point on the elliptic orbit to the foci is constant and equal to the length of the major semi-axis, so $D_1 + D_2 = 2a$ (D_1 and D_2 are the distances to the foci from the point on the ellipse). Since the minor semi-axis is a symmetry axis, its intersection with the elliptic orbit has the property that its distance to both the foci is equal, so $D_1 + D_2 = 2D = 2a$, or:

$$D = a. \tag{381}$$

For the distance D of the intersection of the minor semi-axis to the central particle (the origin), Pythagoras tells us that:

$$D = \sqrt{b^2 + c^2}. \tag{382}$$

Consider now that:

$$\sqrt{b^2 + c^2} = \sqrt{A^2B^2 + \frac{1}{4}A^4 - \frac{1}{2}A^2B^2 + \frac{1}{4}B^4} = \sqrt{\frac{1}{4}A^4 + \frac{1}{2}A^2B^2 + \frac{1}{4}B^4} = a. \tag{383}$$

Thus, the central particle (the origin), must be a focus of the elliptic orbit of the orbital particle. The eccentricity e is defined as the ratio between the distance between the two foci and the lenght of the major semi-axis:

$$e := \frac{2c}{2a} = \frac{c}{a} = \frac{A^2 - B^2}{A^2 + B^2}, \tag{384}$$

which we can use to write:

$$A = \sqrt{\frac{A^2 + B^2}{2} + \frac{A^2 - B^2}{2}} = \sqrt{\frac{A^2 + B^2}{2}(1 + e)} = \sqrt{a(1 + e)} \tag{385}$$

and for B:

$$B = \sqrt{\frac{A^2 + B^2}{2} - \frac{A^2 - B^2}{2}} = \sqrt{\frac{A^2 + B^2}{2}(1 - e)} = \sqrt{a(1 - e)}. \tag{386}$$

Thus, we can finally write:

$$\mathbf{x} = a(e + \cos E) + ia\sqrt{1 - e^2} \sin E \tag{387}$$

and:

$$r = \frac{A^2 + B^2}{2} + \frac{A^2 - B^2}{2} \cos E = a(1 + e \cos E). \tag{388}$$

Proof: see proof 6.5 in *Appendix 1: Proofs*.