Interpreting the Kustaanheimo–Stiefel transform in gravitational dynamics

Prasenjit Saha[*]

Institute for Theoretical Physics, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

ABSTRACT

The Kustaanheimo–Stiefel (KS) transform turns a gravitational two-body problem into a harmonic oscillator, by going to four dimensions. In addition to the mathematical-physics interest, the KS transform has proved very useful in *N*-body simulations, where it helps to handle close encounters. Yet the formalism remains somewhat arcane, with the role of the extra dimension being especially mysterious. This paper shows how the basic transformation can be interpreted as a rotation in three dimensions. For example, if we slew a telescope from zenith to a chosen star in one rotation, we can think of the rotation axis and angle as the KS transform of the star. The non-uniqueness of the rotation axis encodes the extra dimension. This geometrical interpretation becomes evident on writing KS transforms in quaternion form, which also helps to derive concise expressions for regularized equations of motion.

Key words: stellar dynamics – celestial mechanics.

1 INTRODUCTION

The Kustaanheimo–Stiefel (KS) transform is a remarkable relation between the two most important elementary problems in dynamics: under a transformation of coordinates and time, a Kepler problem changes into a harmonic oscillator. Especially noteworthy is that the collision singularity in the Kepler problem is transformed into a regular point. The name comes from the works by Kustaanheimo (1964) and Kustaanheimo & Stiefel (1965), while the book by Stiefel & Scheifele (1971), which is largely devoted to the KS transform and its consequences, is perhaps the best-known source. For a very short summary, see 'regularization' in Binney & Tremaine (2008). An important application of the KS transformation is in numerical orbit integration, where the singularity removal is used to great advantage for simulating dense stellar systems with near collisions (Aarseth & Zare 1974a,b; Jernigan & Porter 1989; Mikkola & Aarseth 1990, 1993). Some recent papers also re-examine the formalism itself (Bartsch 2003; Waldvogel 2006).

In two dimensions there is a much simpler version of the KS transform going back to Levi-Civita (1920). In the Levi-Civita transform, the coordinate plane is read as the complex plane, and the complex square root of the coordinate becomes the transformed coordinate. The geometrical interpretation is clear: the complex phase gets halved. The KS transform is also a kind of square root, but in four dimensions. One wonders how the geometrical interpretation generalizes.

It turns out that slewing a telescope is a convenient geometrical analogy. Suppose the telescope is at zenith and we want to slew it to

a particular star in one rotation. Normally we would simply move along a great circle from the zenith to the star. However, we might prefer a different rotation (to avoid crossing the moon, say). For example, we could choose the mid-point on the above great circle and rotate about it by 180°. In any case, the chosen rotation axis and the rotation angle are effectively the KS transform of the star. This idea of rotation in three dimensions about a non-unique axis generalizes the idea of halving the phase in a complex square root.

This paper attempts to provide some new insight into the KS transform by providing some reformulations and new derivations of known results, and especially to make the geometrical interpretation evident.

2 QUATERNIONS AND ROTATION

Before considering KS theory, it is useful to review a concise algebraic way of specifying rotations in three dimensions, not often used in astrophysics but standard in computer graphics: quaternions.

Quaternions are a generalization of complex numbers. The $\sqrt{-1}$ of complex numbers is replaced by three unit quaternions $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$, such that

$$\boldsymbol{ii} = \boldsymbol{jj} = \boldsymbol{kk} = -1, \qquad \boldsymbol{ijk} = -1. \tag{1}$$

From (1) it follows that $\boldsymbol{ij} = \boldsymbol{k} = -\boldsymbol{ji}$ and so on. In other words, quaternions are like a combination of dot and cross products in vector algebra. (Although historically quaternions came before, having been invented by none other than W. R. Hamilton of Hamilton's equations.)

A general quaternion has the form

$$\boldsymbol{A} = A_0 + A_x \boldsymbol{i} + A_y \boldsymbol{j} + A_z \boldsymbol{k}, \tag{2}$$

[*]E-mail: psaha@physik.uzh.ch

where we will call A_0 the real part. A quaternion with no real part is effectively a vector in three dimensions.

In analogy with complex numbers, we will use the following notation for quaternion conjugates and absolute values:

$$\text{re}[A] \equiv A_0,$$
$$A^* \equiv A_0 - A_1 i - A_2 j - A_3 k,$$
$$A^2 \equiv A^* A = A_0^2 + A_1^2 + A_2^2 + A_3^2. \tag{3}$$

It is easy to see that re $[A^*] = \text{re}[A]$ and $(AB)^* = B^*A^*$, and as a result

$$\text{re}[AB] = \text{re}[B^*A^*] = \text{re}[BA]. \tag{4}$$

Rotation in quaternion notation is beautifully concise. Say we want to rotate a vector r by angle ω about a unit vector n. Using quaternion algebra the rotation is simply

$$R^* r R, \tag{5}$$

where

$$R = \cos \frac{1}{2}\omega + \sin \frac{1}{2}\omega \, n. \tag{6}$$

Unlike the equivalent expression using Euler angles, the expression (5) has no coordinate singularities (or 'gimbal lock') and as a result is numerically more stable, which explains its popularity in computer graphics.

For an arbitrary (i.e. non-unit) quaternion R, the expression (5) amounts to a rotation combined with scalar multiplication.

It is possible to represent quaternions as matrices (though not necessary, even for numerical work). A familiar representation is in terms of Pauli matrices:

$$i = i\sigma_1 \quad j = -i\sigma_2 \quad k = i\sigma_3 \tag{7}$$

or

$$i = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad k = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{8}$$

Pauli matrices are most important as operators on quantum two-state systems (being Hermitian, whereas quaternions are anti-Hermitian). In recent years, the most exciting two-state quantum systems have been Qbits in quantum computing. It turns out that expressions of the type (5) appear in the description of quantum-computing gates (see Mermin 2007, who also provides a derivation of essentially the above three-dimensional rotation formula).

3 THE KUSTAANHEIMO–STIEFEL TRANSFORM

Let

$$q = x i + y j + z k \tag{9}$$

denote a point in space. The KS transform of q is the quaternion

$$Q = Q_0 + Q_x i + Q_y j + Q_z k, \tag{10}$$

the transformation formula being

$$q = Q^* k \, Q. \tag{11}$$

A solution for Q is

$$Q^{\text{I}} = \frac{x i + y j + Z k}{\sqrt{2Z}}, \qquad Z \equiv z + \sqrt{x^2 + y^2 + z^2}, \tag{12}$$

as is easily verified by multiplication, following the quaternion rules. However, Q^{I} is not unique, because changing to

$$Q = (\cos \psi - \sin \psi \, k) \, Q^{\text{I}} \tag{13}$$

leaves equation (11) invariant. Thus ψ behaves like a gauge.

Everything so far is already in the literature. The new result in this paper is that we can readily visualize Q, including its non-uniqueness.

Comparing (11) and (5), it is evident that Q is a rotator that takes the z-axis to q. To visualize Q, let us rewrite q as

$$q = r(\sin\theta\cos\phi \, i + \sin\theta\sin\phi \, j + \cos\theta \, k), \tag{14}$$

where r, θ and ϕ are the usual polar coordinates. Rewriting Q^{I} in the solution (12) and simplifying, we have

$$Q^{\text{I}} = \sqrt{r}\left(\sin\frac{1}{2}\theta\cos\phi \, i + \sin\frac{1}{2}\theta\sin\phi \, j + \cos\frac{1}{2}\theta \, k\right). \tag{15}$$

In other words, the zenith distance of Q^{I} is half-way along the great circle from k to q. From (6) we see the rotation angle ω would be π. Now let us apply the gauge transformation (13) with $\psi = \pi/2$ to Q^{I}. This gives

$$Q^{\text{II}} = \sqrt{r}\left(\cos\frac{1}{2}\theta + \sin\frac{1}{2}\theta\sin\phi \, i - \sin\frac{1}{2}\theta\cos\phi \, j\right). \tag{16}$$

Now the implied rotation is by θ, about an axis perpendicular to both k and q. In general, we can write

$$Q = \cos\psi \, Q^{\text{I}} - \sin\psi \, Q^{\text{II}}, \tag{17}$$

which is to say, Q could be anywhere on the great circle joining Q^{I} and Q^{II}. The telescope-slewing analogy given above is simply a description of the preceding three formulas.

An interesting special case is $\phi = 0$, which gives $q = r(\cos\theta \, k + \sin\theta \, i)$ and $Q^{\text{I}} = \sqrt{r}[\cos(\theta/2) \, k + \sin(\theta/2) \, i]$. Then Q^{I} is effectively the complex square root of q (we need to read k as the real axis and i as the imaginary axis). In other words, the planar case can be reduced to the Levi-Civita transform by a suitable gauge.

Quaternion formulations of the KS transform have been discussed by several authors: Stiefel & Scheifele (1971) mention quaternions but appear to dislike them, while later authors (for example Vivarelli 1994; Waldvogel 2006) are more favourable. The precise definition adopted for the transform varies, but is equivalent to equation (11). That Q represents a rotation and shrinking/stretching of q is also known. Bartsch (2003) specifically notes that the rotation axis is unique in two dimensions but not in three. However, the explicit description of the implied rotations, as above, appears to be new.

4 THE CANONICAL MOMENTUM

So far we have just discussed geometry, but of course the real significance of the KS transform is dynamics, which we now consider. Let

$$p = p_x i + p_y j + p_z k \tag{18}$$

be the canonical momentum conjugate to q. We seek

$$P = P_0 + P_x i + P_y j + P_z k \tag{19}$$

that will be canonically conjugate to Q. Let us write

$$\text{re}[p^* \mathrm{d}q] = \text{re}[p^* \mathrm{d}Q^* k \, Q] + \text{re}[p^* Q^* k \, \mathrm{d}Q]. \tag{20}$$

Using the identity (4), we can rewrite the middle term as re$[p \, (\mathrm{d}Q^* k \, Q)^*]$. Since $p = -p^*$ and $(\mathrm{d}Q^* k \, Q)^* = Q^*(-k)\mathrm{d}Q$ the term becomes becomes re$[p^* Q^* k \, \mathrm{d}Q]$. Thus we have

$$\text{re}[p^* \mathrm{d}q] = 2\text{re}[p^* Q^* k \, \mathrm{d}Q]. \tag{21}$$

Now if we define

$$P = -2k\,Q\,p, \tag{22}$$

we have

$$\mathrm{re}[p^*\mathrm{d}q] = \mathrm{re}[P^*\mathrm{d}Q] \tag{23}$$

which is to say, $p \cdot \mathrm{d}q = P \cdot \mathrm{d}Q$. Provided the Hamiltonian depends on P, Q only through p, q and not on the gauge ψ, the transformation $(P, Q) \rightarrow (p, q)$ is canonical.

To get an explicit expression for p, we multiply (22) on the left by Q^*k, obtaining

$$p = \frac{Q^*kP}{2Q^2}. \tag{24}$$

Note that while we have to be careful about the order of multiplication when i, j, k are involved, real numbers like Q^2 commute with everything. Since p has no real part, $\mathrm{re}[Q^*kP] = 0$ identically. We can think of it as a formal constant of motion resulting from invariance with respect to ψ.

That P (as defined in equation 22, or equivalently) completes a canonical transformation is a standard part of KS theory, but the derivation of the canonical condition using quaternion identities appears to be new.

5 THE TWO-BODY PROBLEM AND THE HARMONIC OSCILLATOR

Let us now write the Kepler Hamiltonian

$$H = \frac{1}{2}p^2 - 1/q \tag{25}$$

in terms of KS variables. Multiplying each of (11) and (22) by its quaternion conjugate, we have

$$q^2 = Q^4, \qquad P^2 = 4p^2Q^2 \tag{26}$$

and substituting these gives

$$H = \frac{1}{8}P^2/Q^2 - 1/Q^2. \tag{27}$$

We now use a device known in Hamiltonian dynamics as a Poincaré time transformation. This involves introducing a fictitious time variable s, whose relation to t we choose to be

$$\mathrm{d}t = Q^2\,\mathrm{d}s. \tag{28}$$

Since Q^2 is the radial distance in the Kepler problem, (28) is in fact Kepler's equation, and s is the eccentric anomaly. In the fictitious time variable s, the equations of motion are given by a new Hamiltonian

$$\Gamma = Q^2(H - E) = \frac{1}{8}P^2 - EQ^2 - 1 \tag{29}$$

with E being the constant initial value of H. The time-transformed Γ Hamiltonian is zero along a trajectory, but its partial derivatives are not zero.

The Hamiltonian Γ is remarkable indeed. For $E < 0$ (bound orbits) it is a harmonic oscillator. Since Q has four components, Γ is like a mass on an isotropic spring in four Euclidean dimensions. Thus the well-known fact that the bound Kepler problem has a dynamical O(4) symmetry. For the unbound case, the symmetry group is different: formally the Lorentz group, but with a physical meaning completely different from special relativity. And – perhaps most importantly – Hamilton's equations for Γ are well behaved even at $Q = 0$ (a collision). This is known as regularization and was the original motivation for KS theory.

The effect of an external force F is simple. From (22) it follows immediately that F will add an extra contribution of $-2k\,Q\,F$ to $\mathrm{d}P/\mathrm{d}t$, which amounts to a contribution of $-2Q^2k\,Q\,F$ to $\mathrm{d}P/\mathrm{d}s$. Provided the external force is non-singular, the equations of motion in s remain regular.

6 REGULARIZING THE THREE-BODY PROBLEM

Application of KS regularization to *N*-body simulations involves expressing the gravitating system either as a tree-like hierarchy of coupled two-body systems (Jernigan & Porter 1989) or as a chain (Mikkola & Aarseth 1990, 1993). The basic idea can be described using the three-body problem with all masses unity. Here again, quaternions enable a concise formulation.

In relative coordinates, the Hamiltonian for three unit gravitating masses can be written (cf. equation 12 in Aarseth & Zare 1974a) as

$$H = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 + p_1 \cdot p_2 - \frac{1}{q_1} - \frac{1}{q_2} \tag{30}$$

plus an additional potential $V(q_1, q_2)$. Here q_1, q_2 expresses the position of the first and second body relative to the zeroth body, while p_1, p_2 express the momenta of the first and second bodies in the barycentric frame. Meanwhile, $V(q_1, q_2)$ expresses the mutual interaction of the first and second bodies, plus any external potential. We can regard $V(q_1, q_2)$ as an external potential, and since we already know how to deal with external forces, we set V aside and concentrate on H.

Now we introduce KS variables $q_1 = Q_1^* k Q_1$ and so on. Defining

$$\Pi \equiv \mathrm{re}\left[(Q_1^*kP_1)^* Q_2^*kP_2\right], \tag{31}$$

we can write $p_1 \cdot p_2$ as $\Pi/(4Q_1^2Q_2^2)$. Applying a Poincaré time transformation

$$\mathrm{d}t = Q_1^2Q_2^2\mathrm{d}s, \qquad \Gamma = Q_1^2Q_2^2(H - E) \tag{32}$$

gives

$$\Gamma = \frac{1}{8}P_1^2Q_1^2 + \frac{1}{8}P_2^2Q_2^2 + \frac{1}{4}\Pi - Q_2^2 - Q_1^2 - EQ_1^2Q_2^2, \tag{33}$$

where E is the value of H. The Γ Hamiltonian has no denominators, and is thus regular for collisions with the zeroth body. (We assume V remains regular, that is to say, the first and second bodies do not collide with each other or any other bodies than the zeroth. In practice, simulations redefine the relative coordinates whenever necessary, according to who is close to whom.)

For the equations of motion, we need derivatives with respect to quaternion components. First we have $\nabla_{Q_1}Q_1^2 = 2Q_1$. Slightly more subtle is $\nabla_{Q_1}\mathrm{re}[Q_1^*A] = A$ if A is independent of Q_1. Using this last identity, together with the definition (31) of Π, we derive

$$\nabla_{P_1}\Pi = -k\,Q_1\,Q_2^*kP_2, \qquad \nabla_{Q_1}\Pi = -k\,P_1\,P_2^*k\,Q_2. \tag{34}$$

Hamilton's equations are then

$$\frac{\mathrm{d}Q_1}{\mathrm{d}s} = \frac{1}{4}Q_1^2P_1 - kP_1P_2^*k\,Q_2,$$

$$\frac{\mathrm{d}P_1}{\mathrm{d}s} = \left(2 + 2EQ_2^2 - \frac{1}{4}P_1^2\right)Q_1 + k\,Q_1\,Q_2^*kP_2 \tag{35}$$

and similarly for P_2, Q_2.

7 DISCUSSION

In dynamical astronomy the KS transformation is profound, but may appear mysterious. This paper attempts to make it less mysterious, and hopefully therefore more useful, by explaining it

in three-dimensional geometric terms. There are several possible directions in which the KS transformation may turn out to be useful.

First, one can imagine new orbit integrators specialized to nearly Keplerian problems. Work on dense stellar systems with near collisions has already been mentioned (for reviews see the books Aarseth 2003; Heggie & Hut 2003). In the planetary regime, which differs from the dense-stellar case in having few bodies but many more orbital times, time transformations reminiscent of (28) used for KS regularization have proved useful for highly eccentric orbits (Mikkola 1997; Emel'yanenko 2002), while some integration algorithms (Mikkola & Tanikawa 1999; Preto & Tremaine 1999) apply the time transformation (28) implicitly. Could the KS transformation itself be exploited here? Fukushima (2005) has some further ideas.

Secondly, it is conceivable that KS variables could simplify perturbation theory. Perturbation theory in classical celestial mechanics (see for example Murray & Dermott 2000) is algebraically frighteningly complicated, basically because the natural variables for the unperturbed and perturbed parts (being the Keplerian action-angles and real-space coordinate) are related through an implicit equation. On the other hand, the action-angles of the KS-transformed Kepler problem are explicitly related to space coordinates – the implicit equation is transferred to the time variable. Could some major simplification be achieved through KS variables? Some progress has been made by Vrbik (2006).

Thirdly, the KS transformation might provide new insight into analogous quantum problem. Bander & Itzykson (1966a,b) derive the symmetry groups of the bound and unbound Coulomb problems. These turn out to be the same four-dimensional symmetries as in KS theory. Is the KS transformation implicit in that work?

ACKNOWLEDGMENTS

REFERENCES

Aarseth S. J., 2003, Gravitational N-Body Simulations. Cambridge Univ. Press, Cambridge

Aarseth S. J., Zare K., 1974a, Celest. Mech., 10, 185

Aarseth S. J., Zare K., 1974b, Celest. Mech., 10, 516

Bander M., Itzykson C., 1966a, Rev. Modern Phys., 38, 330

Bander M., Itzykson C., 1966b, Rev. Modern Phys., 38, 346

Bartsch T., 2003, J. Phys. A: Math. Gen., 36, 6963

Binney J., Tremaine S., 2008, Galactic Dynamics. Princeton Univ. Press, Princeton, NJ

Emel'yanenko V., 2002, Celest. Mech. Dynamical Astron., 84, 331

Fukushima T., 2005, AJ, 129, 2496

Heggie D., Hut P., 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics. Cambridge Univ. Press, Cambridge

Jernigan J. G., Porter D. H., 1989, ApJS, 71, 871

Kustaanheimo P., 1964, in Stiefel E., ed., Mathematische Methoden der Himmelsmechanik und Astronautik Mathematisches Forschungsinstitut Oberwolfach, Berichte 1. Bibliographisches Institut Mannheim, Die Spinordarstellung der energetischen Identitäten der Keplerbewegung, p. 330

Kustaanheimo P., Stiefel E., 1965, J. Reine Angew. Math., 218, 204

Levi-Civita T., 1920, Acta Math., 42, 99

Mermin N. D., 2007, Quantum Computer Science: An Introduction. Cambridge Univ. Press, New York

Mikkola S., 1997, Celest. Mech. Dynamical Astron., 67, 145

Mikkola S., Aarseth S. J., 1990, Celest. Mech. Dynamical Astron., 47, 375

Mikkola S., Aarseth S. J., 1993, Celest. Mech. Dynamical Astron., 57, 439

Mikkola S., Tanikawa K., 1999, Celest. Mech. Dynamical Astron., 74, 287

Murray C. D., Dermott S. F., 2000, Solar System Dynamics. Cambridge Univ. Press, Cambridge

Preto M., Tremaine S., 1999, AJ, 118, 2532

Stiefel E. L., Scheifele G., 1971, Linear and Regular Celestial Mechanics. Springer-Verlag, Berlin

Vivarelli M. D., 1994, Celest. Mech. Dynamical Astron., 60, 291

Vrbik J., 2006, New Astron., 11, 366

Waldvogel J., 2006, Celest. Mech. Dynamical Astron., 95, 201

This paper has been typeset from a TEX/LATEX file prepared by the author.