LINEARIZATION: LAPLACE VS. STIEFEL

ANDRÉ DEPRIT,* ANTONIO ELIPE and SEBASTIÁN FERRER
Grupo de Mecánica Espacial,
Universidad de Zaragoza, 50009 Zaragoza, Spain

(Received 6 July 1992; in final form 8 June, 1993)

Abstract. The method for processing perturbed Keplerian systems known today as the linearization was already known in the XVIIIth century; Laplace seems to be the first to have codified it. We reorganize the classical material around the Theorem of the Moving Frame. Concerning Stiefel's own contribution to the question, on the one hand, we abandon the formalism of Matrix Theory to proceed exclusively in the context of quaternion algebra; on the other hand, we explain how, in the hierarchy of hypercomplex systems, both the KS-transformation and the classical projective decomposition emanate by doubling from the Levi-Civita transformation. We propose three ways of stretching out the projective factoring into four-dimensional coordinate transformations, and offer for each of them a canonical extension into the moment space. One of them is due to Ferrándiz; we prove it to be none other than the extension of Burdet's focal transformation by Liouville's technique. In the course of constructing the other two, we examine the complementarity between two classical methods for transforming Hamiltonian systems, on the one hand, Stiefel's method for raising the dimensions of a system by means of weakly canonical extensions, on the other, Liouville's technique of lowering dimensions through a Reduction induced by ignoration of variables.

Key words: Linearization, KS-transformation, Keplerian problem, Canonical transformations, Hamiltonian systems.

Find what the Sailor has hidden.
VLADIMIR NABOKOV

Introduction

In recognizing that most equations in celestial mechanics are Hamiltonian in nature, Jacobi broke away from traditional procedures when he offered to solve them approximately as perturbations of separable systems. For a long while before him, mathematicians had concentrated their effort on finding the right coordinates and independent variables in order to split the equations into sets of perturbed linear oscillators. Decomposing the position vector $\boldsymbol{x}$ of a mass point into the product of the distance $r = \|\boldsymbol{x}\|$ and the radial direction $\boldsymbol{u} = \boldsymbol{x}/r$ is a preliminary step one is most likely to take in Point Dynamics at the outset of a problem, especially when no symmetry is available to offer an alternative. The pair $(r, \boldsymbol{u})$ makes what, after Ferrándiz, we call the *projective* coordinates of the mass point. In such coordinates, the motion of the particle appears to combine a radial displacement with a rotation of the radial direction on a unit sphere. Changing the independent variable converts the equations of motion in projective coordinates into a linear system; should the

* Permanent address: National Institute of Standards and Technology, Gaithersburg, MD 20899, U. S. A.

Celestial Mechanics and Dynamical Astronomy **58:** 151–201, 1994.

system be Keplerian, the coefficients in the linear system are constant. This artifice converts a perturbed Keplerian system into a set of perturbed harmonic oscillators. The idea behind the scheme is one of the valuable assets in the heritage of celestial mechanics. The technique became standard procedure in the hands of Clairaut, Euler and d'Alembert. It bred a plethora of artifices and stratagems, most of them now obsolete, yet still much alive in the folklore of celestial mechanics. Eventually, Laplace codified the basic algorithms for solving linear differential equations with constant coefficients in Book II of his *Mécanique céleste* published in 1799, the fateful An VII in the Republican Calendar.

But Laplace did nothing toward articulating the techniques for finding transformations that would convert Kepler's problem into a set of harmonic oscillators. On that score, he was a "man in a hurry." He wanted most to pose in the eyes of the new master of the Republic as the French Newton, the one who brought the final solution to the Lunar Theory, a problem that had humbled Newton and his successors, luminaries like Euler and d'Alembert. Owing to Laplace's political opportunism, considering also that his much touted theory of the moon petered out, astronomers of the XIX$^{\text{th}}$ century went on contributing stratagems to linearize equations for individual state functions, leaving to their successors to come yet with more bags of tricks they in turn would claim were better suited to other combinations of variables. The literature on that topic is vast, unwieldy and cumbersome, available today only from historical collections in museum observatories. Take for example Hansen's theory of the "ideal frame." It does wonders for integrating numerically orbits of artificial satellites (Abad *et al.*, 1991; Palacios *et al.*, 1991; Palacios *et al.*, 1992). But Peter Andreas Hansen expressed his ideas in *habilitationsschrift* Latin – the lingo a Schwabian cousin of the Limousin student would have spoken to Pantagruel.* To make matters worse, Hansen did not speak in vectors; they had not yet entered the astronomical language. Amazingly though, at the time Sputnik and Vanguard revived celestial mechanics as an indispensable tool in aerospace research and development, engineers resumed the old habits: an haphazard search for ad hoc linearizations in particular situations. On most occasions, we must confess, they end up reinventing the folklore of celestial mechanics.

We have no intention whatsoever of surveying the theme of linearization in celestial mechanics neither in its sources** nor in its evolution. We confine ourselves to the monograph of Stiefel and Scheifele (1971). In the Preface to their book, these authors make confusing declarations about the novelty of their work. It behooves therefore to remind our readers, if any, that linearization has been a cliché in celestial mechanics throughout the eighteenth century. Laplace (1799, Partie I, Livre 2, Chap. ii and v) summarized in an authoritative synthesis his own

* *Œuvres complètes de Maître François Rabelais. Pantagruel*, vol. VI. Paris, Imprimerie Nationale, 1957.

** For historical references we suggest perusing the splendid reviews in the *Encyklopädie der Mathematische Wissenschaften*, notably that of E. W. Brown, "Theorie des Erdmondes,", Vol. VI, Part 2, Section 14, pp. 667–728 and that of K. F. Sundman, "Theorie der Planeten," Vol. VI, Part 2, Section 15, pp. 729–807.

contributions to the subject and those of his predecessors. Laplace's study has lost none of its relevance.

Thus, in the first section of our paper, solely for the purpose of placing in an historical perspective the research undertaken by Stiefel and his followers, we show how to rehabilitate the Thesaurus of celestial mechanics by bringing to it logic, order and clarity. At the example of Milankovich (1939) and Musen (1954; 1961), we address the issue in the framework of vector geometry; in this way, we spare ourselves many analytical complications and makeshift arguments ascribable to selection of coordinates. The numerous linearized formulas scattered through the literature of the XVIII$^{\text{th}}$ and XIX$^{\text{th}}$ centuries we regroup around a handful of vectorial identities, most of which proceed directly from Darboux's Theorem of the Moving Frame. We start the reconstruction with the exposition made by Laplace. Once given a vectorial formulation, his classical formulas take on symmetric, concise, even elegant forms (Section 1). More importantly, a vectorial treatment makes one realize that linearization as a procedure for solving differential equations pertains to most, if not all, three-dimensional problems in Point Dynamics.

Eduard Stiefel had for the authors of the XVIII$^{\text{th}}$ and XIX$^{\text{th}}$ centuries the tolerance one would expect from an algebraist of renown turned numerical analyst – minimal. He had experimented with linearization while integrating satellite orbits for the European Space Operations Centre; he had found the technique to be markedly stable, a definite advantage over what the specialists had realized so far. Yet, the mathematician he was wanted more than numerical stability in routine circumstances. He was looking for a global code, that is to say, one valid for any set of initial conditions. He heard Kustaanheimo (1964) expounding a scheme based on Pauli's representation of the atom of hydrogen by spinors; he jumped on the idea and joined with Kustaanheimo (Kustaanheimo *et al.*, 1965) in tying linearization with regularization by spinors. By raising the dimensions, he would change Keplerian orbits in three-dimensional Euclidean space into geodesics on a sphere in four dimensions. This gave him more than a handle on a systematic method for linearizing the problem; by completing Kustaanheimo's scheme with a change of independent variable, he would regularize the Keplerian system for its collision orbits. But Kustaanheimo's coordinate transformation raises by one the original dimension of the system. This feature motivated Stiefel (1968) and Scheifele (1968; 1970) to stipulate conditions under which a dimension-raising map in the coordinate space can be extended to the moment space without wiping the Hamiltonian character of the equations provided, of course, that one relaxes the reversibility requirement. Stiefel and Scheifele (1971, Theorem 2, § 31, pp. 189–190) refer still to such extensions as canonical ones. Let it be said right here that this abuse of language is most regrettable, for it has created confusions. To prevent them, it should be enough to speak of extensions of Stiefel type as being *weakly* canonical.

With a criterion in their hands to ascertain that a dimension-raising transformation is weakly canonical, Stiefel and Scheifele went about proposing momenta conjugate to the components of the spinor in the position x. At a crucial point in the construction, they faced a singularly "unwieldy" factorization of polynomials. They chose to skip it in favor of a geometric argument (Stiefel *et al.*, 1971, §43, 270–281). True, a bit of geometry sheds much light onto the nature of the KS-transformation. Yet, the working conditions have changed much over the last twenty years. The symbol processors of today make it relatively easy to cut through algebraic complications and clear a straight path to the hub of the KS-transformation. We avail ourselves of these tools in Section 2 where we review the basic ingredients of the KS-formalism. Free as we are now of computational servitudes, we even put ourselves to the task of resetting the whole KS-theory in terms of quaternions. By the time Stiefel and Scheifele had completed their monograph, they became aware of the close connection between their matrix formalism and the theory of quaternions. It was suggested that they take advantage of it; they reacted to the suggestion in excessive terms (Stiefel *et al.*, 1971, §44, p. 286). Did they really believe that a transfer from matrices to quaternions would lead to "failure or at least to a very unwieldy formalism"? *Chi lo sa*? Stiefel's dire predictions notwithstanding, we accepted the challenge. Did we fail? The reader is our jury. Building the KS-transformation as the emanation of an alternate bilinear form over the algebra of quaternions costs no more in complications than the matrix formalism of Stiefel and Scheifele. Besides, we find rewards in the exercise: theorems are sharpened, some to a significant extent; proofs are shortened; the overall design of Stiefel comes out much enhanced as to its global and intrinsic meaning, not to mention as a collateral a programming style for manipulating quaternions through general purpose Symbol Processors.

The KS-transformation is modelled after the conversion from Cartesian to parabolic coordinates in a plane and its standard canonical extension, a transformation Szebehely* named after Levi-Civita. The link between the two concepts is very close. Velte (1978) sees the KS-transformation as an LC-transformation followed by a rotation in three dimensions. There opened a deeper insight into the issue when Vivarelli (1983) remarked that the KS-transformation is a *doubling* of the LC-transformation. Better still, as one applies meticulously to her papers what Erwin Panowsky ** calls the "postulate of clarification's sake", one unravels from them a general technique for raising two-dimensional transformations to four dimensions and beyond. This possibility is what makes Vivarelli's insight so exciting.

Give up the formalism of the KS-matrix altogether, concentrate on the algebra of quaternions, accept that, in accordance with the multiplication rule of Dickson, quaternions grow by doubling from complex numbers, octonions from quaternions, sedenions from octonions, and so on. Then realize that Dickson's rule is best implemented in software by fashioning hypercomplex numbers after saturated binary trees, that is to say, as pairs of binary trees that are themselves pairs of binary trees, and so on. Now you hold an all-encompassing scheme by which

* *Theory of orbits. The restricted problem of three bodies*, Academic Press, 1967, p. 97

** *Gothic Architecture and Scholasticism*, Latrobe, PA, Archabbey Press, 1951.

operations over hypercomplex numbers of height n spawn operations over hypercomplex numbers of height $n + 1$. That process of induction is of the essence for "doubling." In Section 3.1, we emphasize two points. Firstly: Core operations like addition, scaling, conjugation, multiplication, left conjugation and cross product can all be defined inductively by doubling, and this without recourse to multiplication tables and decomposition into linear combinations of base vectors. Secondly: Inductive doubling translates into efficient schemes for computers to process hypercomplex numbers. Depending on how one interprets the LC-map in the complex plane, one gets distinct ways of doubling it in the algebra of quaternions. Vivarelli's version begets the KS-transformation and its weakly canonical extension, ours in Section 3.3 leads to a standard canonical extension of the three-dimensional projective decomposition. This new kind of doubling provides the pattern matching two canonical extensions of Laplace's projective decomposition we had built some years ago as simpler substitutes for the KS-transformation – and too hastily remanded to oblivion.

About alternatives to the KS-transformation, Stiefel and Scheifele (1971, p. 288) issued a warning little short of an injunction: "the authors are convinced that the search for other transformations [. . .] is not very promising." To many readers their omen conveyed the impression that the KS-technique is unique in achieving jointly regularization, linearization and dimension-raising for three-dimensional Keplerian systems. The facts disallow the claim. Kustaanheimo, Stiefel and Scheifele never mention the decisive step Fock (1935; 1936) had taken in that direction thirty-five years earlier, not even in their brief allusion to the same step taken, but independently, by Moser (1970). We have room here for the latest entry in the competition, the BF-transformation due to Burdet (1969) for the coordinate part and completed by Ferrándiz (1986a; 1986b; 1987; 1988) for the moment part. We supplement it with a transformation of our own, the DEF-transformation, which we claim achieves equally well all the objectives of the KS-transformation – linearization, regularization and canonicity – although, we are inclined to believe, in a simpler and more intuitive way (Section 4.1). Admittedly the construction involves heavy algebraic manipulations, no more however than is the case with the KS- or the BF-transformation. Besides, we pass the chore to the Symbol Processor. Yet, lest we create misunderstandings we hasten to emphasize the obvious, that a Symbol Processor is no more than a mathematical accountant. However efficient its accounting methods are, it cannot relieve users from the responsibility of creating simplifications toward clamping final results in their most significant form.

As one deals with weakly canonical transformations, one cannot help but feel that the constructions have been left unfinished, that there might be, in the final analysis, a way of supplementing the original variables with enough coordinates and momenta to make the maps both reversible and canonical. After all, the Fundamental Theorem of Stiefel and Scheifele is meant to deal with dimension-raising transformations that cannot be made, for intrinsic reasons, into extensions of a coordinate transformation in lower dimensions. This, according to Stiefel (1971,

§44, p. 287) is the case for the KS-transformation. Is it also true for the DEF-transformation? We do not answer the question.

While grappling with this issue, we modified the moment segment in the DEF-transformation and came with yet another weakly canonical extension of the projective point-transformation, namely the D-transformation (Section 4.3). This time, however, we found a pair of invariants x_0 and X_0 that we could bind into a couple of conjugate coordinate and momentum. We appended x_0 and X_0 to the position $\boldsymbol{x}$ and its conjugate momentum $\boldsymbol{X}$, we embedded the D-transformation (from an eight-dimensional to a six-dimensional domain) into a genuinely canonical transformation from an eight-dimensional to an eight-dimensional domain. To boot, we found that the D-transformation so completed can be generated in the tritest manner as the gauge-free homogeneous extension of its point-transformation. We conclude the paper in the same vein with Section 4.4. After discussing briefly the contributions of Burdet and Ferrándiz, we show how to extend their 8-to-6 weakly canonical map into an eight-dimensional canonical transformation in the classical sense.

1. Linearization: synthesis in a vector perspective

Laplace's exposition is limited primarily to Keplerian systems and their perturbations; at times, though, it reaches for gradients of a force function. For our part, we plan on pursuing the classical line as far as possible without any restriction. We start thus with the fundamental equation of Point Dynamics $\ddot{\boldsymbol{x}} = \boldsymbol{F}$ where $\boldsymbol{F}$ represents the force (per unit of mass) applied to a particle. We find convenient to handle the equation as a system of two differential equations of the first order. Moreover we introduce the angular momentum (per unit of mass)

$$\boldsymbol{G} = \boldsymbol{x} \times \dot{\boldsymbol{x}}. \tag{1}$$

As a result we can replace the fundamental equation of Point Dynamics by the overabundant system

$$\dot{\boldsymbol{x}} = \boldsymbol{X}, \qquad \dot{\boldsymbol{X}} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{X}, t), \qquad \dot{\boldsymbol{G}} = \boldsymbol{x} \times \boldsymbol{F} \tag{2}$$

1.1. In cylindrical coordinates

Let S be a frame fixed in space made of three orthonormal vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$. We decompose the position vector $\boldsymbol{x}$ into the sum

$$\boldsymbol{x} = (\boldsymbol{x} \cdot \boldsymbol{e}_3)\, \boldsymbol{e}_3 + (\boldsymbol{e}_3 \times \boldsymbol{x}) \times \boldsymbol{e}_3$$

of its components respectively parallel and perpendicular to $\boldsymbol{e}_3$; its projection onto the reference plane $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ we factor into the product

$$(\boldsymbol{e}_3 \times \boldsymbol{x}) \times \boldsymbol{e}_3 = \rho \boldsymbol{m} \quad \text{with} \quad \rho > 0 \quad \text{and} \quad \|\boldsymbol{m}\| = 1.$$

We also set $z = \boldsymbol{x} \cdot \boldsymbol{e}_3$ and $\boldsymbol{e} = \boldsymbol{e}_3 \times \boldsymbol{m}$. In addition, we introduce the longitude of the particle in the plane $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ as the angle λ such that

$$\boldsymbol{m} = \boldsymbol{e}_1 \cos\lambda + \boldsymbol{e}_2 \sin\lambda \quad \text{with} \quad 0 \leq \lambda < 2\pi,$$

and the component of the angular momentum

$$\Lambda = \boldsymbol{G} \cdot \boldsymbol{e}_3 = \rho^2 \dot{\lambda}.$$

It is then deduced readily that (2) is equivalent to the differential system

$$\begin{cases} \ddot{z} = \boldsymbol{F} \cdot \boldsymbol{e}_3, \\ \ddot{\rho} = \dfrac{\Lambda^2}{\rho^3} + (\boldsymbol{F} \cdot \boldsymbol{m}), \\ \dot{\Lambda} = \rho(\boldsymbol{F} \cdot \boldsymbol{e}), \\ \dot{\lambda} = \dfrac{\Lambda}{\rho^2}. \end{cases} \tag{3}$$

These are the equations of motion in cylindrical coordinates.

In order to *linearize* them Laplace replaced the variables z and ρ by the quantities σ and ζ, and also the independent variable t by an angle f_0, such that

$$\sigma = 1/\rho, \quad \zeta = z/\rho, \quad \text{and} \quad \rho^2\, df_0 = \Lambda\, dt.$$

For the sake of conciseness, we set $\boldsymbol{F}_0 = \rho^3 \boldsymbol{F}/\Lambda^2$; we note that the vector $\boldsymbol{F}_0$ is dimensionless. We now claim that the equations in (3) are equivalent to the system

$$\begin{aligned} &\frac{d^2\zeta}{df_0^2} + \zeta = -\zeta\,(\boldsymbol{F}_0 \cdot \boldsymbol{m}) - \frac{d\zeta}{df_0}\,(\boldsymbol{F}_0 \cdot \boldsymbol{e}) + (\boldsymbol{F}_0 \cdot \boldsymbol{e}_3), \\ &\frac{d^2\sigma}{df_0^2} + \sigma = -\sigma\,(\boldsymbol{F}_0 \cdot \boldsymbol{m}) - \frac{d\sigma}{df_0}\,(\boldsymbol{F}_0 \cdot \boldsymbol{e}), \\ &\qquad \frac{d\Lambda}{df_0} = \Lambda(\boldsymbol{F}_0 \cdot \boldsymbol{e}). \end{aligned} \tag{4}$$

We spare readers (if any) an account of the train of deductions that led to recast Laplace's equations in so symmetric a format; rather we invite them to check the formulas, preferably with a general purpose Symbol Processor. In the case when $\boldsymbol{F}$ is the gradient of a function $\mathcal{U}(\boldsymbol{x}, t)$ with respect to $\boldsymbol{x}$, the system corresponds to the set of equations labelled (K) in Laplace (1799, Livre II, Ch. ii, §15). Brown (1896, p. 19) reconstructed the set but in its scalar form and only when $\boldsymbol{F}$ is the gradient of a force function, so he could spare himself the chore of identifying meaningful components among the terms in the right hand members of (4).

Lunar theory is what motivated Clairaut, Euler, d'Alembert and Laplace to linearize system (3). In that problem, the force may be split as a sum of the type

$$\boldsymbol{F} = -\frac{\mu}{\rho^2}\boldsymbol{m} + \epsilon \boldsymbol{P}$$

where ϵ is a token parameter (that is to say, a parameter with no functionality other than that of ranking terms on an asymptotic scale). The perturbation being omitted from the model, equations in (4) break in two pieces: a pair of linear equations of order two with constant coefficients

$$\left(\frac{d^2}{df_0^2} + 1\right)\sigma = \frac{\mu}{\Lambda^2} \quad \text{and} \quad \left(\frac{d^2}{df_0^2} + 1\right)\zeta = 0,$$

plus a quadrature for the longitude, owing to the fact that, in the absence of a perturbation, the polar component Λ of the angular momentum is an integral.

1.2. In Spherical Coordinates

Cylindrical coordinates achieve linearization when the principal component of the force presents an axial symmetry. In the absence of an axial symmetry, linearization is best pursued in the orbital frame of reference.

We start with the projective decomposition of the position as the product

$$\boldsymbol{x} = r\boldsymbol{u} \quad \text{with} \quad r > 0 \quad \text{and} \quad \|\boldsymbol{u}\| = 1.$$

When turning our attention to the time derivative of $\boldsymbol{x}$ so decomposed, we call on the following statement:

LEMMA 1. – *Given $\boldsymbol{a}(t) \neq 0$ and the projective decomposition*

$$\boldsymbol{a} = \alpha\boldsymbol{b} \quad \text{with} \quad \alpha > 0 \quad \text{and} \quad \|\boldsymbol{b}\| = 1,$$

there follows that

$$\dot\alpha = \boldsymbol{b}\cdot\dot{\boldsymbol{a}} \quad \text{and} \quad \dot{\boldsymbol{b}} = \alpha^{-2}(\boldsymbol{a}\times\dot{\boldsymbol{a}})\times\boldsymbol{b}.$$

In application of this rule; one finds that

$$\dot r = \boldsymbol{u}\cdot\boldsymbol{X} \quad \text{and} \quad \dot{\boldsymbol{u}} = \frac{\boldsymbol{G}}{r^2}\times\boldsymbol{u}.$$

Likewise, from the projective decomposition

$$\boldsymbol{G} = G\boldsymbol{n} \quad \text{with} \quad G > 0 \quad \text{and} \quad \|\boldsymbol{n}\| = 1,$$

one obtains at once that

$$\dot G = r\,(\boldsymbol{v}\cdot\boldsymbol{F}) \quad \text{and} \quad \dot{\boldsymbol{n}} = -\frac{r}{G}\,(\boldsymbol{n}\cdot\boldsymbol{F})\,\boldsymbol{v},$$

$\boldsymbol{v}$ being the unit vector $\boldsymbol{v} = \boldsymbol{n} \times \boldsymbol{u}$. The vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{n}$ are respectively the radial, transverse and normal directions; they define a frame $\mathcal{O}$ to be referred to as the *orbital* frame.

It now becomes clear that the original Newtonian equations are equivalent to the system of dimension 12 made of three vector equations

$$\dot{\boldsymbol{u}} = \boldsymbol{\omega} \times \boldsymbol{u}, \quad \dot{\boldsymbol{v}} = \boldsymbol{\omega} \times \boldsymbol{v}, \quad \dot{\boldsymbol{n}} = \boldsymbol{\omega} \times \boldsymbol{n}$$

to account for the rotation of the orbital frame $\mathcal{O}$ at the angular velocity (see Darboux (1915))

$$\boldsymbol{\omega} = \frac{G}{r^2}\,\boldsymbol{n} + \frac{r}{G}\,(\boldsymbol{n} \cdot \boldsymbol{F})\boldsymbol{u},$$

and of the two scalar equations

$$\dot{G} = r\,(\boldsymbol{v} \cdot \boldsymbol{F}), \quad \ddot{r} = \frac{G^2}{r^3} + (\boldsymbol{u} \cdot \boldsymbol{F})$$

to account for the motion in the radial direction. The unknowns $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{n}$ satisfy the invariant orthonormality relations

$$||\boldsymbol{u}|| = ||\boldsymbol{v}|| = ||\boldsymbol{n}|| = 1, \quad \text{and} \quad \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{w} \cdot \boldsymbol{u} = 0.$$

We now propose to replace the independent variable t and the distance r respectively by the dimensionless variable f (call it an angle) and the quantity q such that

$$r^2\,df = G\,dt \quad \text{and} \quad q = 1/r.$$

Accordingly, we rescale the force and the angular velocity to make the vectors

$$\boldsymbol{F}^* = \frac{r^3}{G^2}\boldsymbol{F} \quad \text{and} \quad \boldsymbol{\omega}^* = \frac{r^2}{G}\,\boldsymbol{\omega} = \boldsymbol{n} + (\boldsymbol{n} \cdot \boldsymbol{F}^*)\boldsymbol{u}$$

dimensionless. After a few elementary manipulations we find that the orbital equations take on the following form:

$$\begin{cases} \dfrac{d}{df}\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \\ \boldsymbol{n} \end{pmatrix} = \boldsymbol{\omega}^* \times \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \\ \boldsymbol{n} \end{pmatrix}, \\[2ex] \dfrac{dG}{df} = G\,(\boldsymbol{v} \cdot \boldsymbol{F}^*), \\[2ex] \dfrac{d^2 q}{df^2} + q = -q\,(\boldsymbol{u} \cdot \boldsymbol{F}^*) - \dfrac{dq}{df}\,(\boldsymbol{v} \cdot \boldsymbol{F}^*). \end{cases} \tag{5}$$

Let us assume for a while that $\boldsymbol{F}$ belongs to the class of vector fields for which $\boldsymbol{F} \cdot \boldsymbol{n} = 0$. Incidentally, the class includes non central forces like drag. Under that assumption, the angular velocity $\boldsymbol{\omega}^*$ is identical to $\boldsymbol{n}$, hence the direction $\boldsymbol{n}$ is an integral. Moreover,

$$\frac{d^2\boldsymbol{u}}{df^2} = \frac{d}{df}(\boldsymbol{n} \times \boldsymbol{u}) = \frac{d\boldsymbol{v}}{df} = \boldsymbol{n} \times \boldsymbol{v} = -\boldsymbol{u}.$$

The equation restates the obvious. A system made of a particle acted upon by a force permanently locked within the orbital plane is invariant for the rotation group SO(3), which invariance entails that the radial direction is in free rotation about the origin. Of course, Laplace and his predecessors did not acknowledge the symmetry group as such; for them, rotational invariance meant that the longitude λ and the latitude β are solutions of the homogeneous linear system

$$\left(\frac{d^2}{df^2} + 1\right) \begin{pmatrix} \cos\lambda\cos\beta \\ \sin\lambda\cos\beta \\ \sin\beta \end{pmatrix} = 0. \tag{6}$$

These formulas, most certainly for pure Keplerian systems, have been known for a long time, so long a time indeed that they keep falling into oblivion out of which, owing to their elementary character, they are periodically recalled to be greeted – sometimes not without irony – as surprises.

1.3. In Orbital Coordinates

Even when the force has a normal component, the equation for the radial direction is still linear in its principal part (i.e. the part obtained by equating the perturbation to 0), but it is no longer homogeneous. Indeed, from the modified equations for the rotation of the orbital frame, we deduce easily that

$$\frac{d^2\boldsymbol{u}}{df^2} = \frac{d}{df}(\boldsymbol{\omega}^* \times \boldsymbol{u}) = \frac{d\boldsymbol{v}}{df} = \boldsymbol{\omega}^* \times \boldsymbol{v} = -\boldsymbol{u} + (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,\boldsymbol{u}\times\boldsymbol{v},$$

hence the second order equation for the radial direction,

$$\left(\frac{d^2}{df^2} + 1\right)\boldsymbol{u} = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,\boldsymbol{n}. \tag{7}$$

We need at this point the classical elements that specify the orbit and the position of the particle on the orbit:

(i) the inclination as the angle I such that

$$e_3 \cdot \boldsymbol{n} = \cos I \quad \text{with} \quad 0 \le I \le \pi;$$

it goes without saying that we do not consider the singular cases when $I = 0$ or $I = \pi$.

(ii) under the assumption that I is neither 0 nor π, the ascending node $\boldsymbol{l}$ as the direction such that

$$e_3 \times \boldsymbol{n} = \boldsymbol{l}\sin I \quad \text{and} \quad \|\boldsymbol{l}\| = 1;$$

(iii) the longitude ν of the ascending node as defined by the conditions

$$\boldsymbol{l} = \boldsymbol{e}_1 \cos\nu + \boldsymbol{e}_2 \sin\nu \quad \text{and} \quad 0 \le \nu < 2\pi;$$

(iv) the argument of latitude θ such that

$$\boldsymbol{l} = \boldsymbol{u}\cos\theta - \boldsymbol{v}\sin\theta \quad \text{and} \quad 0 \le \theta < 2\pi. \tag{8}$$

Thanks to these definitions, composing equation (7) in a dot product with the base vector $\boldsymbol{e}_3$ yields the Latitude Equation

$$\left(\frac{d^2}{df^2} + 1\right) \sin\beta = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,(\boldsymbol{n}\cdot\boldsymbol{e}_3) = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,\cos I;$$

doing the same with $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ produces the Longitude Equations

$$\begin{cases} \left(\dfrac{d^2}{df^2} + 1\right) \cos\lambda\cos\beta = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,(\boldsymbol{n}\cdot\boldsymbol{e}_1) = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,\sin\nu\sin I, \\ \left(\dfrac{d^2}{df^2} + 1\right) \sin\lambda\cos\beta = (\boldsymbol{n}\cdot\boldsymbol{F}^*)\,(\boldsymbol{n}\cdot\boldsymbol{e}_2) = -(\boldsymbol{n}\cdot\boldsymbol{F}^*)\,\cos\nu\sin I. \end{cases}$$

Longitude and Latitude Equations are commonly regarded in celestial mechanics as long time acquisitions, so much so that they are usually quoted without attribution [see, e.g., Brown and Shook (1933, § 1.28, pp. 29–30)].

As an illustration, take the main problem in the theory of artificial satellites. The force function being there

$$\mathcal{U} = \frac{\mu}{r}\left[1 - J_2\left(\frac{\alpha}{r}\right)^2 P_2(\boldsymbol{u}\cdot\boldsymbol{e}_3)\right],$$

its gradient splits into the sum

$$\boldsymbol{F} = -\frac{\mu}{r^2}\boldsymbol{u}\left[1 + \frac{3}{2}J_2\left(\frac{\alpha}{r}\right)^2\left(1 - 5(\boldsymbol{u}\cdot\boldsymbol{e}_3)^2\right)\right] - 3J_2\frac{\mu}{r^2}\left(\frac{\alpha}{r}\right)^2(\boldsymbol{u}\cdot\boldsymbol{e}_3)\boldsymbol{e}_3.$$

Therefore

$$\boldsymbol{n}\cdot\boldsymbol{F}^* = -3J_2\frac{\mu\alpha^2}{G^2 r}\sin\beta\cos I,$$

hence the Equation of Latitude

$$\left(\frac{d^2}{df^2} + 1\right) \sin\beta = -3J_2\frac{\mu\alpha^2}{G^2 r}\sin\beta\cos^2 I$$

recovered by Jezewski (1983, Formula 26, p. 351).

Finally, we return to the case of forces without a normal component. In that special situation, on the one hand, the node $\boldsymbol{l}$ is a fixed direction; on the other, according to (8),

$$\boldsymbol{u} = \boldsymbol{l}\cos\theta + \boldsymbol{n}\times\boldsymbol{l}\sin\theta.$$

Therefore, in view of (7),

$$\left(\frac{d^2}{df^2}+1\right)\cos\theta = (\boldsymbol{n}\cdot\boldsymbol{F}^*)(\boldsymbol{n}\cdot\boldsymbol{l}) = 0,$$
$$\left(\frac{d^2}{df^2}+1\right)\sin\theta = (\boldsymbol{n}\cdot\boldsymbol{F}^*)(\boldsymbol{n}\cdot(\boldsymbol{n}\times\boldsymbol{l})) = 0.$$

Brown paraphrasing Laplace refers to these identities as the Equations of the True Longitude. Like in the case of the Longitude and Latitude Equations, a vectorial treatment gets the Equations of the True Longitude to stand for what they are, that is, trivial consequences of Darboux's Theorem of the Moving Frame. Invoking that commonplace theorem spared ourselves the kind of *unwieldy* and *cumbersome* treatment – these are the words of Stiefel and Scheifele – that linearization of Point Dynamics received lately.

Before we leave this review, we must mention another line of linearization schemes. It opened at the turn of this century with Bohlin (1911). Breaking away from the Laplacian tradition of using a true anomaly as the independent variable, Bohlin observed that, with t' such that $r\,dt' = \alpha\,dt$ in place of t, the equations of motion for the problem of two bodies are of the form:

$$\left(\frac{d^2}{dt'^2}-\frac{2\mathcal{H}}{\alpha^2}\right)\boldsymbol{x}+\frac{\boldsymbol{A}}{\alpha^2} = 0.$$

The functions $\mathcal{H}$ and $\boldsymbol{A}$ stand respectively for the energy and the Laplace-Hamilton vector $\boldsymbol{A} = \boldsymbol{X}\times\boldsymbol{G} - \mu\boldsymbol{u}$ whereas α is a length scale introduced to give t' the physical dimension of a time. Wintner (1947, § 259 and p. 423) who knew Bohlin's paper tied Bohlin's linearization with the Levi-Civita transformation. Some twenty years later, Bohlin's idea resurfaced in the articles of Burdet (1967; 1968; 1969), and from there provided Stiefel with a guide toward other types of linearization.

2. The KS-transformation: a quaternion version

Stiefel and Scheifele designed the KS-transformation by tracing off the LC-transformation in two dimensions. In several places, they have been confronted with overdemanding algebraic manipulations like factoring polynomials of degree six in eight variables, which tasks they either set aside or circumvented by resorting to geometric constructions. We take a hint given by Stiefel and Scheifele (1971, §44, p. 286) for reconstructing the KS-transformation and its canonical extension in terms of quaternions. We insist on treating the KS-transformation for itself regardless of its possible application to Keplerian systems, in other words, merely as an exhibit in the gallery of canonical maps over an eight-dimensional phase space. Here and there we take the opportunity for sharpening several theorems stated by Stiefel. As for the situations leading to extensive calculations, we do not dodge them; rather, we commit them to a Symbol Processor – we use *Mathematica* of Wolfram (1991).

2.1. THE ALGEBRA OF QUATERNIONS

Let us begin with a few words about notations in the set $\mathcal{Q}$ of quaternions. Any element q of $\mathcal{Q}$ is here represented as a list $\{u_0, u_1, u_2, u_3\}$ of four real numbers, with u_0 standing for the *scalar* part of u while u_1, u_2 and u_3 form its *vector* part $u^\natural$. We say that u is a *pure* vector when $u = u^\natural$. The standard base is made of the four quaternions $e_0 = (1,0,0,0)$, $e_1 = (0,1,0,0)$, $e_2 = (0,0,1,0)$ and $e_3 = (0,0,0,1)$. We denote by $\overline{u}$ the conjugate of the quaternion u.

With the symbolic Processor we are using, we found it more expedient to define the product of quaternion p by quaternion q as the operation that produces a quaternion r whose scalar and vector parts are respectively

$$r_0 = p_0 q_0 - p^\natural \cdot q^\natural \quad \text{and} \quad r^\natural = p_0 q^\natural + q_0 p^\natural + p^\natural \times q^\natural.$$

Multiplication of quaternions is not commutative. Yet, for any u, v and w in $\mathcal{Q}$,

$$(uvw - \overline{w}v\overline{u})_0 = 2(uvw)_0 - 2v_0(uw)_0,$$
$$(uvw - \overline{w}v\overline{u})^\natural = 2v_0(uw)^\natural,$$

hence this weak form of commutativity which will resurface later in the most basic properties of the KS-operators:

LEMMA 2. – *If v is a pure vector, then $uvw - \overline{w}v\overline{u}$ is a pure scalar, and*

$$(uvw - \overline{w}v\overline{u})_0 = 2(uvw)_0.$$

The expression $u \cdot v$ designates the scalar product of the quaternions u and v; for the norm of u we use the notation $\|u\|$. The lemma below states that the operator $v \longrightarrow uv : \mathcal{Q} \longrightarrow \mathcal{Q}$ of multiplication to the left by u transposes through the dot product into multiplication to the left by $\overline{u}$; symmetrically, multiplication to the right by u transposes into multiplication to the right by $\overline{u}$.

LEMMA 3. – *For any u, v, and w in $\mathcal{Q}$, $u\,v \cdot w = v \cdot \overline{u}\,w = u \cdot w\,\overline{v}$.*

Proof. – On the one hand,

$$\begin{aligned} uv \cdot w &= \tfrac{1}{2}((uv)\overline{w} + w\overline{(uv)})_0 \\ &= \tfrac{1}{2}(uv\overline{w} + w\overline{v}\,\overline{u})_0 \\ &= \tfrac{1}{2}(u\overline{(w\overline{v})} + (w\overline{v})\overline{u})_0 \\ &= u \cdot w\overline{v}. \end{aligned}$$

On the other hand,

$$uv \cdot w = \overline{(uv)} \cdot \overline{w} = \overline{v}\,\overline{u} \cdot \overline{w},$$

which, by virtue of what has just been proved, implies that

$$uv \cdot w = \overline{v} \cdot \overline{w}u = v \cdot \overline{u}w. \blacksquare$$

The monograph of Stiefel and Scheifele contains the rudiments of a notion of cross product in $\mathbf{R}^4$; Vivarelli (1988) expanded them somewhat. We do not have the place here to examine the interaction between the structure of exterior algebra and that of quaternion algebra over $\mathbf{R}^4$. We shall content ourselves with defining the cross product of two quaternions as the bilinear operator

$$(u, v) \longrightarrow u \wedge v = \tfrac{1}{2}(v\overline{u} - u\overline{v}) : \mathcal{Q} \times \mathcal{Q} \longrightarrow \mathcal{Q}.$$

The definition implies that the quaternion $u \wedge v$ is a pure vector; moreover

$$(u \wedge v)^\natural = u_0 v^\natural - v_0 u^\natural + u^\natural \times v^\natural.$$

In particular, $(e_0 \wedge u)^\natural = u^\natural$ for any $u \in \mathcal{Q}$. Note that all cross products $e_i \wedge e_j$ are zero save

$$\begin{aligned} e_0 \wedge e_i &= e_i \quad \text{for} \quad 1 \leq i \leq 3, \\ e_1 \wedge e_2 &= -e_2 \wedge e_1 = e_3, \\ e_2 \wedge e_3 &= -e_3 \wedge e_2 = e_1, \\ e_3 \wedge e_1 &= -e_1 \wedge e_3 = e_2. \end{aligned}$$

The cross product is skew-symmetric, i.e., $u \wedge v = -v \wedge u$. It is not associative – we beg to disagree on this point with Vivarelli (1988, p. 361, Eq. 7): for example, $(e_1 \wedge e_0) \wedge e_2 = -e_3$ whereas $e_1 \wedge (e_0 \wedge e_2) = e_3$. In general,

$$\begin{aligned} ((u \wedge v) \wedge w)^\natural = (u^\natural \times v^\natural) \times w^\natural + u_0(v^\natural \times w^\natural - w_0 v^\natural) \\ - w_0(u^\natural \times v^\natural - v_0 u^\natural) + v_0(w^\natural \times u^\natural), \end{aligned}$$

hence the rule of weak associativity

$$\begin{aligned} ((u \wedge v) \wedge w)^\natural - (u \wedge (v \wedge w))^\natural = \\ (u \cdot v)w^\natural - (v \cdot w)u^\natural - 2v_0(u \wedge w)^\natural. \end{aligned} \tag{9}$$

When restricted to the subspace of pure vectors, the cross product of quaternions is identical to the cross product of vectors; more precisely

$$(u \wedge v)^\natural = u^\natural \times v^\natural \quad \text{when} \quad u_0 = v_0 = 0.$$

This fact, however, should not invite to manipulate cross products of quaternions like cross products of vectors. On the contrary, there is so much incompatibility between the two operations that one is well advised to use different notations for them. For instance, the formula

$$((u \wedge v) \wedge w)^\natural + ((v \wedge w) \wedge u)^\natural + ((w \wedge u) \wedge v)^\natural = u_0(v^\natural \times w^\natural) + v_0(w^\natural \times u^\natural) + w_0(u^\natural \times v^\natural)$$

tells that Jacobi's identity is satisfied in $\mathcal{Q}$ only if all three operands are pure vectors. In the same vein, from the formula

$$(u \wedge v) \cdot w = u^\natural \times v^\natural \cdot w^\natural + u_0(v \cdot w) - v_0(u \cdot w), \tag{10}$$

one deduces readily the identity

$$(u \wedge v) \cdot w - u \cdot (v \wedge w) = u_0(v \cdot w) - 2v_0(w \cdot u) + w_0(u \cdot v)$$

which manifests that dot and cross products are interchangeable only if their operands are pure vectors. Nevertheless, some rules of vector algebra apply unconditionally to quaternions too. As an illustration, start from the identity

$$\begin{aligned}(u \wedge v) \cdot (w \wedge t) &= (u \cdot w)(v \cdot t) - (u \cdot t)(v \cdot w) \\ &\quad + u_0(v^\natural \cdot (w^\natural \times t^\natural)) - v_0(w^\natural \cdot (t^\natural \times u^\natural)) \\ &\quad + w_0(t^\natural \cdot (u^\natural \times v^\natural)) - t_0(u^\natural \cdot (v^\natural \times w^\natural))\end{aligned}$$

for any u, v, w and t in $\mathcal{Q}$. Set therein $t = v$; the result is a formula

$$(u \wedge v) \cdot (w \wedge v) = \|v\|^2 (u \cdot w) - (u \cdot v)(v \cdot w), \tag{11}$$

valid not only for three-dimensional vectors but also for quaternions. In the particular case where $w = u$, (11) is the Lagrange identity

$$\|u \wedge v\|^2 = \|u\|^2 \|v\|^2 - (u \cdot v)^2$$

mentioned by Stiefel and Scheifele (1970, § 43, p. 278).

We conclude these comments about the cross product of quaternions with a generalization of what Vivarelli (1988, p. 363, Eq. 19) calls the "w-relation."

THEOREM 1. – *For any u and v in $\mathcal{Q}$, the relation*

$$((u \wedge v) \wedge \overline{w})^\natural = (w \wedge (\overline{u} \wedge \overline{v}))^\natural$$

holds true for any w that is a linear combination of u and v.

Proof. – For any w in $\mathcal{Q}$,

$$((u \wedge v) \wedge \overline{w})^\natural - (w \wedge (\overline{u} \wedge \overline{v}))^\natural = 2u_0(w^\natural \times v^\natural) + 2v_0(u^\natural \times w^\natural) + 2w_0(v^\natural \times u^\natural).$$

In particular,

$$(u \wedge v) \wedge \overline{u} = u \wedge (\overline{u} \wedge \overline{v}) \quad \text{and} \quad (u \wedge v) \wedge \overline{v} = v \wedge (\overline{u} \wedge \overline{v});$$

hence, for any α and β in $\mathbf{R}$,

$$(u \wedge v) \wedge (\overline{\alpha u + \beta v}) = (\alpha u + \beta v) \wedge (\overline{u} \wedge \overline{v}). \blacksquare$$

2.2. THE KS-OPERATORS AND THE KS-ALTERNATE FORM

To every quaternion u, we associate two linear operators

$$L_u : v \longrightarrow L_u(v) = v e_1 \overline{u} : \mathcal{Q} \longrightarrow \mathcal{Q},$$
$$M_u : v \longrightarrow M_u(v) = v u \overline{e_1} : \mathcal{Q} \longrightarrow \mathcal{Q}.$$

For $u = 0$, $L_0 = M_0 = 0$. Otherwise, the operators are regular since $\det(L_u) = \det(M_u) = \|u\|^4$; moreover, on account of the identity

$$L_u(M_u(v)) = M_u(v) e_1 \overline{u} = v u \overline{e_1}\, e_1 \overline{u} = \|u\|^2 v,$$

there follows that $M_u/\|u\|$ is the inverse of $L_u/\|u\|$. By virtue of Lemma 3,

$$\|L_u(v)\|^2 = \|M_u(v)\|^2 = \|u\|^2\, \|v\|^2,$$

which means that the operators $L_u/\|u\|$ and $M_u/\|u\|$ are orthogonal. There also follows from the same lemma 3 that, for any v and w in $\mathcal{Q}$, $L_u(v) \cdot w = v \cdot M_u(w)$, or that L_u and M_u are transposes of one another.

Alongside the operator L_u, we define the bilinear form

$$J : (u, v) \longrightarrow J(u, v) = u \cdot (v e_1) : \mathcal{Q} \times \mathcal{Q} \longrightarrow \mathcal{R}.$$

The connection between J and the operator L is manifest: $J(u, v)$ is the scalar component of the quaternion $L_u(v)$.

The bilinear form J is non degenerate: since

$$J(e_0, v) = -v_1, \quad J(e_1, v) = v_0, \quad J(e_2, v) = v_3, \quad J(e_3, v) = -v_2,$$

if $J(u, v) = 0$ for any u in $\mathcal{Q}$, then $v = 0$. It is reflexive: by virtue of Lemma 3,

$$J(u, u) = (\overline{u}\, u) \cdot e_1 = \|u\|^2\, (e_0 \cdot e_1) = 0.$$

It is skew-symmetric: evidently, $J(u, v) = -J(v, u)$. For being non degenerate, reflexive and skew-symmetric, J is what algebraists like Dieudonné (1955, p. 12) call an *alternate* form.

THEOREM 2. – *Given two linearly independent quaternions u and v, let* P *be the vector subspace in $\mathcal{Q}$ spanned by u and v. If $J(u, v) = 0$, then $J(x, y) = 0$ for any x and any y in* P.

Proof. – Any element of P is, by assumption, a linear combination of u and v. On the other hand,

$$J(\alpha u + \beta v, \alpha' u + \beta' v) = (\alpha \beta' - \alpha' \beta) J(u, v). \blacksquare$$

Of a vector subspace P of dimension 2 for which the restriction of J to $P \times P$ vanishes identically, algebraists say that it is *totally isotropic* relative to J; Stiefel and Scheifele (1971, p. 273–275) call it a *Levi-Civita plane.*

All coordinate planes $\{e_i, e_j\}$ in $\mathcal{Q}$, i.e., the planes spanned by the base vectors e_i and e_j, are totally isotropic relative to J save the planes $\{e_0, e_1\}$ and $\{e_2, e_3\}$. There is a reason for that:

LEMMA 4. – *Given $u \neq 0$ in $\mathcal{Q}$, a quaternion v is orthogonal to u, equal in norm to u and such that $J(u,v)=0$ if and only if it is of the form*

$$v = u(e_2 \cos\beta + e_3 \sin\beta) \tag{12}$$

Proof. – Since u is $\neq 0$, for any v there is a w in $\mathcal{Q}$ such that $v = uw$. With v so factored, we find that

$$J(v,u) = w_1 \|u\|^2, \quad v \cdot u = w_0 \|u\|^2, \quad \text{and} \quad \|v\|^2 = \|w\|^2 \|u\|^2,$$

and deduce from these equations that v satisfies the three conditions mentioned in the theorem if and only if $w_0 = w_1 = 0$ and $w_2^2 + w_3^2 = 1$. ∎

THEOREM 3. – *A vector subspace* P *of dimension two in $\mathcal{Q}$ is totally isotropic relative to J if and only it has a base consisting of a quaternion u and a quaternion v such that*

$$v = u(e_2 \cos\beta + e_3 \sin\beta) \quad \text{with} \quad 0 \le \beta < \pi.$$

Proof. – That the condition is sufficient follows from Theorem 2. But it is also necessary. Indeed P admits an orthonormal base (u,v), and if P is totally isotropic, then $J(u,v)=0$ and so by virtue of Lemma 4, v must be of the form (12). ∎

Lemma 4 supplies a procedure for constructing all the planes totally isotropic relative to J that contain a given $u \neq 0$. For example, the planes through e_0 that are totally isotropic are those spanned by e_0 and $e_2 \cos\beta + e_3 \sin\beta$; this one-parameter family includes the coordinate planes $\{e_0, e_2\}$ for $\beta = 0$ and $\{e_0, e_3\}$ for $\beta = \pi/2$.

As we shall now see with the next theorem, planes that are totally isotropic relative to J are characterized by the way they affect the behavior of the operators L_u.

THEOREM 4 (Reciprocity of the KS-operators). – *Concerning a vector subspace* P *of dimension two in $\mathcal{Q}$, the following properties are equivalent:*

a) P *is totally isotropic relative to J;*

b) there is a base (u,v) in P *such that $L_u(v) = L_v(u)$;*

c) there is a base (u,v) in P *such that*

$$(u \cdot u)L_v(v) - 2(u \cdot v)L_u(v) + (v \cdot v)L_u(u) = 0.$$

Proof. – That a) $\Longleftrightarrow$ b) follows immediately from the weak commutativity of the product of quaternions. Indeed, e_1 being a pure vector, there comes by virtue of Lemma 2 that

$$L_u(v) - L_v(u) = v e_1 \overline{u} - u\, e_1 \overline{v} = 2(v e_1 \overline{u} \cdot e_0) e_0 = 2J(u,v) e_0. \tag{13}$$

The proof that a) $\Longleftrightarrow$ c) follows the same pattern. First, one should realize that

$$\begin{aligned}(v\cdot v)L_u(u)+(u\cdot u)L_v(v) &= (v\overline{v})(ue_1\overline{u})+(u\overline{u})(ve_1\overline{v})\\ &= (u\overline{v})(ve_1\overline{u})+(\overline{v}u)(ue_1\overline{v});\end{aligned}$$

then, by reason of the weak commutativity (Lemma 2), one obtains that

$$\begin{aligned}(v\cdot v)L_u(u)+(u\cdot u)L_v(v) &= (u\overline{v})(ve_1\overline{u})+(v\overline{u})[ve_1\overline{u}+2(ue_1\overline{v}\cdot e_0)e_0]\\ &= 2(u\cdot v)L_u(v)+2J(v,u)v\overline{u},\end{aligned}$$

hence the identity

$$\begin{aligned}(u\cdot u)L_v(v)-2(u\cdot v)\,L_u(v)+(v\cdot v)\,L_u(u) &=\\ 2J(v,u)\left(u\wedge v+(u\cdot v)e_0\right).\end{aligned} \tag{14}$$

Now, for any $u\neq 0$, the relation $u\wedge v+(u\cdot v)e_0=0$ is satisfied only for $v=0$. Therefore, in the formula (14), the left hand member is zero if and only $J(u,v)=0$. ■

COROLLARY 4.1. – *Let* P *be a vector subspace of dimension two in* $\mathcal{Q}$. *If* P *is totally isotropic relative to* J, *then*

$$L_u(v)=L_v(u)\quad\text{and}\quad (u\cdot u)L_v(v)-2(u\cdot v)L_u(v)+(v\cdot v)L_u(u)=0$$

for any u *and* v *in* P.

The theorem of reciprocity is stronger than Theorems 1 and 2 in Stiefel and Scheifele (1970, § 9, p. 15) which, in fact, correspond to the above corollary.

THEOREM 5. – *For any* u *and* v *in* $\mathcal{Q}$,

a) $(u\wedge v)\cdot L_u(u)=(u\cdot u)\,J(v,u)$;

b) $(u\wedge v)\cdot L_u(v)=(u\cdot v)\,J(v,u)$.

Proof. – By definition of the cross product,

$$(u\wedge v)\cdot L_u(u)=(v\overline{u}-(u\cdot v)e_0)\cdot L_u(u)=(v\overline{u})\cdot L_u(u).$$

Then, by virtue of Lemma 3,

$$(v\overline{u})\cdot L_u(u)=(v\overline{u})\cdot(ue_1\overline{u})=v\cdot(ue_1\overline{u}u)=(u\cdot u)(v\cdot ue_1)=(u\cdot u)\,J(v,u).$$

This completes the proof of part a). As for part c), still by application of Lemma 3, we get that

$$(v\overline{u})\cdot(ve_1\overline{u})=(v\cdot v)(\overline{u}\cdot(e_1\overline{u}))=(u\cdot u)(v\cdot v)(e_0\cdot e_1)=0.$$

Therefore we conclude that

$$(u\wedge v)\cdot L_u(v)=-(u\cdot v)J(u,v)=(u\cdot v)J(v,u).\ ■$$

2.3. The KS-point-transformation

Given $\alpha > 0$ in $\mathbf{R}$, the KS-transformation κ_α is the mapping

$$\kappa_\alpha : v \longrightarrow x = \frac{1}{\alpha} L_v(v) : \mathcal{Q} \longrightarrow \mathcal{Q}, \tag{15}$$

or, explicitly, the mapping defined by the equations

$$\begin{cases} \alpha x_0 = 0, \\ \alpha x_1 = v_1^2 - v_2^2 - v_3^2 + v_0^2, \\ \alpha x_2 = 2(v_1 v_2 + v_3 v_0), \\ \alpha x_3 = 2(v_1 v_3 - v_2 v_0). \end{cases} \tag{16}$$

We have introduced the parameter α to maintain a simple correspondence in dimensions between the quaternions x and v, which correspondence we achieve by attributing the dimension of x to α so that v will likewise be of the same dimension as x. Keeping homogeneity in dimensions is advantageous on two counts. (i) Symbolic calculations demand to be constantly on the alert for mistakes. Checking – by sight or by program – that a formula is correct in dimensions is the fastest way to detect gross errors. (ii) Homogeneity in dimensions means invariance with respect to a group of similitude; exploiting this symmetry often produces interesting results. [See, e.g., Meyer (1984) or Deprit and Williams (1991).]

For the sake of short notations, we do not mention the parameter when referring to a KS-transformation.

Clearly, κ is a homogeneous quadratic transformation, that is, $\kappa(\beta u) = \beta^2 \kappa(u)$ for any $\beta \in \mathbf{R}$; in particular, $\kappa(-u) = kappa(u)$. On the other hand, $\|\kappa(u)\| = \|u\|^2/\alpha$ since

$$\alpha^2 \|\kappa(u)\|^2 = ((ue_1\,\overline{u})(\overline{ue_1\,\overline{u}}))_0 = (ue_1(\overline{u}\,u)\overline{e_1}\,\overline{u})_0 = \|u\|^2\,(u(e_1\overline{e_1})\overline{u})_0 = \|u\|^4.$$

There follows in particular that $\kappa(u) = 0$ if and only if $u = 0$. Moreover, for any $\boldsymbol{x} \neq 0$ in $\mathbf{R}^3$, both quaternions

$$\begin{aligned} y_+ &= \sqrt{\frac{\alpha}{2}(r + x_1)}\;\left[e_1 + \frac{1}{r + x_1}(x_2 e_2 + x_3 e_3)\right] \\ y_- &= \sqrt{\frac{\alpha}{2}(r - x_1)}\;\left[e_2 + \frac{1}{r - x_1}(x_2 e_1 - x_3 e_0)\right] \end{aligned}$$

with $r = \|\boldsymbol{x}\|$ are mapped by κ onto $\boldsymbol{x}$; in other words, κ is a surjection of $\mathcal{Q}$ onto $\mathbf{R}^3$.

Let $\kappa^{-1}(\boldsymbol{x})$ designate the set of all $u \in \mathcal{Q}$ such that $\kappa(u) = \boldsymbol{x}$.

THEOREM 6 (Fibration). – *Given* $\boldsymbol{x}$ *in* $\mathbf{R}^3$ *and* u *in* $\mathcal{Q}$ *such that* $\boldsymbol{x} \neq 0$ *and* $\kappa(u)^\natural = \boldsymbol{x}$, *a) the set of quaternions v such that $\kappa(v) = \boldsymbol{x}$ is the great circle*

$$\Gamma_+(u) = \{u(e_0 \cos\phi + e_1 \sin\phi) \,:\, 0 \leq \phi < 2\pi\};$$

b) the set of quaternions v such that $\kappa(v) = -\boldsymbol{x}$ is the great circle

$$\Gamma_-(u) = \{u(e_2 \cos\phi + e_3 \sin\phi) \,:\, 0 \leq \phi < 2\pi\};$$

Proof. – Indeed, the quaternion $e_0 \cos\phi + e_1 \sin\phi$ represents a rotation with axis $(e_1)^\natural$ and amplitude 2ϕ, hence

$$\begin{aligned}\kappa(u(e_0 \cos\phi + e_1 \sin\phi)) \\ &= u\,(e_0 \cos\phi + e_1 \sin\phi)\, e_1\, \overline{(e_0 \cos\phi + e_1 \sin\phi)}\, \overline{u} \\ &= u e_1 \overline{u} = \kappa(u).\end{aligned}$$

Therefore $\Gamma_+(u) \subseteq \kappa^{-1}(\boldsymbol{x})$. Conversely, the equality $\kappa(u) = \kappa(v)$ implies that

$$(\overline{u}\,v) e_1 (\overline{v}\,u) = \|u\|^4 e_1.$$

But $\alpha\|\boldsymbol{x}\| = \|u\|^2 = \|v\|^2$, hence the quaternion $(\overline{u}\,v)/\|u\|^2$ is unitary. Therefore, according to the preceding relation, it represents a rotation with axis $e_1^\natural$, which means that it is of the form

$$(\overline{u}\,v)/\|u\|^2 = (e_0 \cos\phi + e_1 \sin\phi).$$

There follows that $v = u(e_0 \cos\phi + e_1 \sin\phi)$, or that $\kappa^{-1}(\boldsymbol{x}) \subseteq \Gamma_+(u)$.

A straightforward calculation analogous to the one made in part a) shows that $\kappa(u(e_2 \cos\phi + e_3 \sin\phi)) = -\boldsymbol{x}$. Conversely, the relation $\kappa(v) = -\kappa(u)$ implies that

$$(\overline{u}\,v) e_1 (\overline{v}\,u) = -\|u\|^4 e_1.$$

It means that the quaternion $(\overline{u}\,v)/\|u\|^2$ represents a rotation. We may decompose it into the product $\mathcal{R}(\phi, \mathcal{R}(\pi, e_2^\natural)(e_1^\natural)) \circ \mathcal{R}(\pi, e_2^\natural)$. Yet, according to Rodrigues's Theorem of the Product of Rotations

$$\mathcal{R}(\phi, \mathcal{R}(\pi, e_2^\natural)(e_1^\natural)) \circ \mathcal{R}(\pi, e_2^\natural) = \mathcal{R}(\pi, e_2^\natural) \circ \mathcal{R}(\phi, e_1^\natural).$$

Therefore, the quaternion $(\overline{u}\,v)/\|u\|^2$ is of the form

$$(\overline{u}\,v)/\|u\|^2 = (e_0 \cos\phi + e_1 \sin\phi)(e_0 \cos\tfrac{1}{2}\pi + e_2 \sin\tfrac{1}{2}\pi)$$

which means that $v = u(e_2 \cos\phi + e_3 \sin\phi)$ or that $v \in \Gamma_-(u)$. ∎

Completing a remark made by Velte (1978), we signal the identities

$$\kappa(\xi e_1 + \eta e_2) = \kappa(\xi e_0 + \eta e_3) = (0, (\xi^2 - \eta^2)/\alpha, 2\xi\eta/\alpha, 0)\ ,$$

$$\kappa(\xi e_1 + \eta e_3) = \kappa(\xi e_0 - \eta e_2) = (0, (\xi^2 - \eta^2)/\alpha, 0, 2\xi\eta/\alpha)$$

because they show that the restriction of the KS-transformation to any of the coordinate planes $\{e_0, e_2\}$, $\{e_0, e_3\}$, $\{e_1, e_2\}$ and $\{e_1, e_3\}$ is an LC-transformation from parabolic to Cartesian coordinates. At the outset the KS-transformation was conceived as an extension to higher dimensions of the LC-transformation $\{e_1, e_2\} \longrightarrow \mathbf{R}^2$. Eventually Stiefel and Scheifele (1971, § 43, Theorem 5) came to realize that such relationship holds in many more planes than just the coordinate plane $\{e_1, e_2\}$:

THEOREM 7. – *The restriction of the KS-transformation to any plane that is totally isotropic relative to J is an LC-transformation.*

Proof. –A plane P, we have already remarked after Theorem 4, is totally isotropic if it is spanned by a quaternion $u \neq 0$ and a quaternion v of the form given in (12). The pair (u, v) makes an orthogonal base of P. Observe, though, that the pair $(\kappa(u), \kappa(v))$ does not make a base of $\kappa(P)$ since $\kappa(v) = -\kappa(u)$. To make such a base, take the quaternion $w = (u+v)/\sqrt{2}$ in P. It can be shown readily with a Symbol Processor that

$$\| \kappa(w) \| = \| \kappa(v) \| \quad \text{and} \quad \kappa(w) \cdot \kappa(u) = 0.$$

There follows that, for any ξ and η in $\mathbf{R}$,

$$\kappa(\xi u + \eta v) = (\xi^2 - \eta^2)\kappa(u) + 2\xi\eta\kappa(w),$$

which proves that κ restricted to P is an LC transformation. ∎

Velte (1978) examined the inverse problem: Given a plane P in $\mathbf{R}^3$, an orthogonal base $(\boldsymbol{f}_1, \boldsymbol{f}_2)$ in P, and a LC-transformation

$$\lambda : \xi \boldsymbol{f}_1 + \eta \boldsymbol{f}_2 \longrightarrow \boldsymbol{x} = \frac{\xi^2 - \eta^2}{\alpha} \boldsymbol{f}_1 + \frac{2\xi\eta}{\alpha} \boldsymbol{f}_2 : \mathrm{P} \longrightarrow \mathrm{P}$$

find the quaternions u such that

$$\kappa(u) = \lambda(\xi \boldsymbol{f}_1 + \eta \boldsymbol{f}_2). \tag{17}$$

To solve the problem we introduce the vector $\boldsymbol{f}_3 = \boldsymbol{f}_1 \times \boldsymbol{f}_2$ to make an orthonormal base in $\mathbf{R}^3$, and we designate by f_i the pure vector whose vector component is $\boldsymbol{f}_i$. In those terms, λ is readily seen to be the restriction to P of the mapping

$$\check{\lambda} : v \longrightarrow \check{\lambda}(v) = (v f_1 \overline{v})/\alpha : \mathcal{Q} \longrightarrow \mathcal{Q}.$$

Now let q be one of the two unitary quaternions representing the rotation mapping e_i onto f_i for $i = 1, 2, 3$. From the fact that $f_1 = q e_1 \overline{q}$ we deduce at once that

$$\check{\lambda}(v) = (vq) e_1 (\overline{vq})/\alpha = \kappa(vq).$$

By virtue of Theorem 6, any u such that $\kappa(u) = \kappa(vq)$ is of the form

$$u = vq(e_0 \cos\phi + e_1 \sin\phi).$$

In particular, the roots of equation (17) are the quaternions

$$(\xi f_1 + \eta f_2) q (e_0 \cos\phi + e_1 \sin\phi) = q(\xi e_1 + \eta e_2)(e_0 \cos\phi + e_1 \sin\phi).$$

In Velte (1978) only the solution for which $\phi = 0$ is mentioned.

2.4. AN EXTENSION OF THE KS-POINT-TRANSFORMATION

Stiefel and Scheifele undertook to extend canonically the coordinate transformation (15) into a mapping

$$(v, V) \longrightarrow (x, X) : \mathcal{Q} \times \mathcal{Q} \longrightarrow \mathcal{Q} \times \mathcal{Q}.$$

They built their extension in an *ad hoc* fashion by seeking to maintain an analogy between the three-dimensional KS-transformation and the two-dimensional LC-transformation (Stiefel *et al.*, 1971, §38, p. 234). We prefer the reasoning of Kurcheeva (1977). On account of the definition in (15),

$$\begin{aligned}\alpha\, dx &= (dv)e_1\overline{v} + ve_1\overline{(dv)} = (dv)e_1\overline{v} + \overline{(dv)\overline{e_1}\,\overline{v}}\\ &= (dv)e_1\overline{v} - \overline{(dv)e_1\,\overline{v}} = L_v(dv) - \overline{L_v(dv)},\end{aligned}$$

hence the fundamental formula

$$\alpha\, dx = 2\left[L_v(dv) - J(v, dv)e_0\right].$$

Kurcheeva asks to build the one-form

$$\alpha X \cdot dx = 2\left[X \cdot L_v(dv) - J(v, dv)(X \cdot e_0)\right]$$

We observe that

$$X \cdot L_v(dv) = M_v(X) \cdot dv,$$

and that

$$J(v, dv)(X \cdot e_0) = 0 \quad \text{when} \quad X_0 = 0.$$

This being the case, Kurcheeva chooses

$$V = 2M_v(X)/\alpha \tag{18}$$

to ensure that the one-form becomes the homogeneous identity

$$X_1\, dx_1 + X_2\, dx_2 + X_3\, dx_3 = V \cdot dv.$$

In explicit terms, we find that the new momenta are the functions

$$\begin{aligned}V_0 &= 2\,(X_1v_0 + X_0v_1 - X_3v_2 + X_2v_3)/\alpha,\\ V_1 &= -2\,(X_0v_0 - X_1v_1 - X_2v_2 - X_3v_3)/\alpha,\\ V_2 &= -2\,(X_3v_0 - X_2v_1 + X_1v_2 + X_0v_3)/\alpha,\\ V_3 &= 2\,(X_2v_0 + X_3v_1 + X_0v_2 - X_1v_3)/\alpha\end{aligned}$$

in 'mixed' variables – new coordinates and old momenta. By inversion of (18), we recover the extension made by Stiefel and Scheifele:

$$V \longrightarrow X = \frac{1}{2r} L_v(V) : \mathcal{Q} \longrightarrow \mathcal{Q}, \tag{19}$$

where $\alpha r = \alpha\sqrt{x \cdot x} = v \cdot v$. As a precaution against possible confusions with the corresponding, but not identical, formulas in Stiefel and Scheifele (1971, §38, Equation 31) we develop the expression in (19):

$$X_0 = \tfrac{1}{2}(\ \ V_0 v_1 - V_1 v_0 - V_2 v_3 + V_3 v_2)/r,$$

$$X_1 = \tfrac{1}{2}(\ \ V_0 v_0 + V_1 v_1 - V_2 v_2 - V_3 v_3)/r,$$

$$X_2 = \tfrac{1}{2}(\ \ V_0 v_3 + V_1 v_2 + V_2 v_1 + V_3 v_0)/r,$$

$$X_3 = \tfrac{1}{2}(-V_0 v_2 + V_1 v_3 - V_2 v_0 + V_3 v_1)/r.$$

It should be observed for further reference that

$$X_0 = \tfrac{1}{2} J(v, V)/r. \tag{20}$$

We find it convenient, in view of the applications, to group various facts about basic physical quantities when they are expressed in KS-variables.

THEOREM 8. – *For any x and any X in $\mathcal{Q}$,*

$$x \cdot X = x^{\natural} \cdot X^{\natural} = \tfrac{1}{2}(v \cdot V), \tag{21}$$

$$\|X\|^2 = \frac{\alpha}{4r}\, \|V\|^2, \tag{22}$$

$$x^{\natural} \times X^{\natural} = \frac{1}{2}\Big[v \wedge V + \frac{1}{r} x^{\natural} J(v, V)\Big]. \tag{23}$$

Proof. – By straightforward calculation using the Symbol Processor. ∎

COROLLARY 8.1. *For any x and any X in $\mathcal{Q}$,*

a) $\|X^{\natural}\|^2 = \frac{1}{4}\Big[\frac{\alpha}{r}\,\|V\|^2 - \frac{1}{r^2}\,(J(v, V))^2\Big];$

b) $\|x^{\natural} \times X^{\natural}\|^2 = \frac{1}{4}\,\Big[\|\, v \wedge V\,\|^2 - (J(v, V))^2\Big];$

c) $(v \wedge V) \cdot x^{\natural} = -r\, J(v, V);$

d) $(v \wedge V) \cdot X^{\natural} = -\frac{1}{r}(x^{\natural} \cdot X^{\natural}) J(v, V).$

Proof. – Part a) follows readily from (22) and (20), part b) from (22) and the Lagrange identity for cross product of quaternions mentioned in Section 2.1. As for parts c) and d), they are immediate consequences of (23). ■

For the benefit of clarity, we denote by $\{f\,;\,g\}$ the Poisson bracket of f and g. In the present context,

$$\{f\,;\,g\} = \sum_{0\le i\le 3}\left(\frac{\partial f}{\partial v_i}\frac{\partial g}{\partial V_i} - \frac{\partial f}{\partial V_i}\frac{\partial g}{\partial v_i}\right).$$

With the Symbol Processor, one evaluates readily all the Poisson brackets entering the theory of Stiefel and Scheifele. We list first those which are the quickest to evaluate:

$$\{x_i\,;\,x_j\} = 0 \quad \text{for} \quad 0\le i,j\le 3,$$
$$\{x_0\,;\,X_i\} = \{x_i\,;\,X_0\} = 0 \quad \text{for} \quad 0\le i\le 3,$$
$$\{x_i\,;\,X_j\} = \delta_{i,j} \quad \text{for} \quad 1\le i\le j\le 3.$$

Thereafter we engage the Symbol Processor into the arduous task of proving the following statement:

THEOREM 9. – *Let P be the matrix formed of the Poisson brackets $\{X_i\,;\,X_j\}$ for $0\le i,j\le 3$, and let $P(w)$ denote the quaternion image of a quaternion w by the matrix P. Then*

$$P(w) = \frac{1}{2r^3}J(v,V)wx \tag{24}$$

holds uniformly in v and V by virtue of the KS-transformation.

Replacing w by each of the standard base quaternions in (24) yields the matrix of Poisson brackets

$$P = \frac{J(v,V)}{2r^3}\begin{pmatrix} 0 & -x_1 & -x_2 & -x_3 \\ x_1 & 0 & x_3 & -x_2 \\ x_2 & -x_3 & 0 & x_1 \\ x_3 & x_2 & -x_1 & 0 \end{pmatrix}. \tag{25}$$

Incidentally, we must inform the reader that the corresponding formula in the monograph of Stiefel and Scheifele (1971, §38, Equation 56) is in error. It should read:

$$(\{p_i,p_j\})_{1\le i,j\le 3} = \frac{1}{2|\overline{x}|^6}(\overline{x}_1\overline{p}_4 - \overline{x}_2\overline{p}_3 + \overline{x}_3\overline{p}_2 - \overline{x}_4\overline{p}_1)\begin{pmatrix} 0 & x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{pmatrix}.$$

We now have in our hands the tools we need to apply the KS-transformation to the Hamiltonian

$$\mathcal{H} = \frac{1}{2}(X_1^2 + X_2^2 + X_3^2) - \frac{\mu}{r} \tag{26}$$

of a Keplerian system. On account of the identity in part a) of Corollary 8.1,

$$\mathcal{H} = \frac{1}{8}\left[\frac{\alpha}{r}(V \cdot V) - \frac{J(v,V)^2}{r^2}\right] - \frac{\mu}{r}. \tag{27}$$

To the independent variable t we substitute a variable t' such that

$$2r\,dt' = \alpha\,dt.$$

By virtue of our convention on the dimension of the parameter α, t' is of the same dimension as t. With the new independent variable, the Cartan one-form governing the motion on the manifold $\mathcal{H} = h$ becomes

$$V \cdot dv - (\mathcal{H} - h)\,dt = V \cdot dv - \frac{2r}{\alpha}(\mathcal{H} - h)\,dt',$$

which form prompts us to substitute for $\mathcal{H}$ the Hamiltonian

$$\mathcal{K} = \frac{2r}{\alpha}(\mathcal{H} - h) + \frac{2\mu}{\alpha} = \frac{1}{4}(V \cdot V) - 2h\frac{r}{\alpha} - \frac{J(v,V)^2}{4\alpha r}. \tag{28}$$

We do this for the purpose of establishing a one-to-one correspondence between, on the one hand, the orbits (parametrized by t) on the manifold $\mathcal{H} = h$ which are solutions of the canonical equations

$$\begin{aligned}
\frac{dv}{dt} &= \{v\,;\,\mathcal{H}\} = \frac{\alpha}{4r}V - \frac{J(v,V)}{4r^2}\,\nabla_V J(v,V),\\
\frac{dV}{dt} &= \{V\,;\,\mathcal{H}\} = \frac{1}{\alpha r}\left(2h - \frac{J(v,V)^2}{4r^2}\right)v + \frac{J(v,V)}{4r^2}\,\nabla_v J(v,V)
\end{aligned}$$

and, on the other, the orbits (parametrized by t') on the manifold $\mathcal{K} = 2\mu/\alpha$ which are solutions of the canonical equations

$$\begin{aligned}
\frac{dv}{dt'} &= \{v\,;\,\mathcal{K}\} = \frac{1}{2}V - \frac{J(v,V)}{2\alpha r}\,\nabla_V J(v,V),\\
\frac{dV}{dt'} &= \{V\,;\,\mathcal{K}\} = \left(\frac{4h}{\alpha^2} - \frac{J(v,V)^2}{2\alpha^2 r^2}\right)v + \frac{J(v,V)}{2\alpha r}\,\nabla_v J(v,V).
\end{aligned} \tag{29}$$

We finally come to the constraint that characterizes the KS-transformation, namely that it be restricted to the quadric

$$\mathcal{J} = \{(v,V)\;:\;J(v,V) = 0\}.$$

The restriction is legitimate from a dynamical standpoint. Indeed, by virtue of (20) and (25), THEOREM 10. – *The relations*

$$\{J(v,V)\,;\,x_i\} = \{J(v,V)\,;\,X_i\} = 0 \quad \text{for} \quad 0 \leq i \leq 3$$

hold uniformly in v *and* V *when the variables* x *and* X *are replaced by their expressions in (15) and (19).*

The theorem says that $\mathcal{J}$ is an integral manifold for any system represented by a Hamiltonian F that is a function of x and X. On the invariant manifold $\mathcal{J}$,

(i) X_0 vanishes identically, and the KS-transformation may then be regarded as a dimension-raising map

$$(v,V) \longrightarrow (x_1,x_2,x_3,X_1,X_2,X_3) : \mathcal{J} \longrightarrow \mathbf{R}^3 \times \mathbf{R}^3;$$

(ii) The conditions for the KS-transformation to be weakly canonical are satisfied since, with $J(v,V) = 0$,

$$\{x_i\,;\,x_j\} = \{X_i\,;\,X_j\} = 0 \quad \text{and} \quad \{x_i\,;\,X_j\} = \delta_{i,j} \quad \text{for} \quad 1 \leq i \leq j \leq 3;$$

(iii) The KS-transformation linearizes and regularizes the Keplerian system. Indeed, on the invariant manifold $\mathcal{J}$,

$$\left\{\frac{dv}{dt'}, \mathcal{K}\right\}\Big|_{J=0} = \frac{2h}{\alpha^2}\,v,$$

hence the equations in (29) take on the simple form

$$\left(\frac{d^2}{dt'^2} - \frac{2h}{\alpha^2}\right) v = 0.$$

3. Doublings of an LC-transformation

Giacaglia (1970, p. 20) may have been the first to have observed that the coordinate segment in the KS-transformation is of the form $x = u\,\phi(u)$ where u is a quaternion and ϕ a skew involution for quaternions. But he does it in passing without attaching much importance to that remark. Later Vivarelli made the same observation, but aligned it with the fact that the LC-transformation

$$(\xi,\eta) \longrightarrow (\xi^2 - \eta^2, 2\xi\eta) : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$$

of the two-dimensional real plane onto itself can also be expressed as the mapping

$$\xi + i\eta \longrightarrow (\xi + i\eta) * (\xi + i\eta) \;:\; \mathbf{C} \longrightarrow \mathbf{C}$$

of the complex line onto itself, that is to say, as a product of hypercomplex numbers. Considering that complex numbers make the first generation of hypercomplex numbers produced from the reals by doubling, and that quaternions are the second generation, she came to realize that the KS-transformation derives from the LC-transformation by doubling too. At the next turn of the recursive crank, Lambert and Kibler (1988) produce the Hurwitz transformations over the alternative algebra of Cayley numbers.

Vivarelli's approach broke new ground; it is important to us because, as we shall see, it leads to a new type of extension of the LC-transformation, in fact a pattern for the canonical transformations to be detailed in Section 4.

3.1. The Doubling Technique

We review the theory of hypercomplex structures for the twofold purpose of (a) emphasizing the *doubling* technique, and (b) indicating how the technique ushers in an efficient way of representing hypercomplex structures in a Symbol Processor. By taking advantage of it, we can afford to ignore the usual method of coding hypercomplex numbers as arrays of their components. Much is gained in brevity and transparency since manipulating these algebraic objects as binary trees spares the chore of repeatedly addressing multiplication tables.

At the outset is a non empty set A. (For the applications we have in mind, A is the field $\mathbf{R}$ of real numbers.) On top of A is built a sequence

$$\mathrm{A}^{(1)}, \mathrm{A}^{(2)}, \mathrm{A}^{(3)}, \ldots, \mathrm{A}^{(n)}, \mathrm{A}^{(n+1)}, \ldots \tag{30}$$

of sets by induction over $n \geq 1$: $\mathrm{A}^{(1)}$ is the set of pairs (z_l, z_r) of elements of A, that is, the product set $\mathrm{A} \times \mathrm{A}$, $\mathrm{A}^{(2)}$ the product set $\mathrm{A}^{(1)} \times \mathrm{A}^{(1)}$ and, in general, $\mathrm{A}^{(n+1)}$ the product set $\mathrm{A}^{(n)} \times \mathrm{A}^{(n)}$, or to say it in other words, $\mathrm{A}^{(n+1)}$ is obtained by *doubling* its predecessor in the sequence. We speak of the elements of $\mathrm{A}^{(n)}$ as the *binary trees* over A with height n. With a Symbol Processor at hand, one finds it natural to represent binary trees as lists of two elements, each one being in turn a pair of binary trees; the height of the tree materializes as the *depth* of its representative list.

Transfer from parallel array formalism to inductive list processing is accomplished through a constructor τ for converting sequences of 2^n elements in A into binary trees of height n:

$$\tau(z) = \begin{cases} z & \text{for } z \in \mathrm{A}, \\ \Big(\tau(z_0, \ldots, z_{2^{(n-1)}-1}), \tau(z_{2^{(n-1)}}, \ldots, z_{2^n-1})\Big) & \text{for } z \in \mathrm{A}^{(n)}. \end{cases}$$

When applied to the rows of the identity matrix of dimension 2^n, the constructor τ produces the standard base in the vector space $\mathrm{A}^{(n)}$. For instance, τ produces the trees

$$e_0 = ((1,0),(0,0)), \qquad e_1 = ((0,1),(0,0)),$$
$$e_2 = ((0,0),(1,0)), \qquad e_3 = ((0,0),(0,1))$$

as the standard base in the hypercomplex system $\mathrm{A}^{(2)}$ of quaternions . The inverse operation τ^{-1} converts binary trees into sequences; to say it in typographic terms, τ^{-1} strikes all parentheses off a binary tree save the first and the last. Current

Symbol Processors like MACSYMA, MAPLE, MATHEMATICA, REDUCE perform the operation through a command like `flatten`.

Two elements of any $\mathrm{A}^{(n)}$ must be singled out for their properties with respect to addition and multiplication: the zero element which is the binary tree

$$0^{(n)} = \begin{cases} (0,0) & \text{for } n = 1, \\ (0^{(n-1)}, 0^{(n-1)}) & \text{for } n > 1, \end{cases}$$

and the unit element

$$1^{(n)} = \begin{cases} (1,0) & \text{for } n = 1, \\ (1^{(n-1)}, 0^{(n-1)}) & \text{for } n > 1. \end{cases}$$

Let it be assumed now that A is equipped with an addition. By inductive doubling, the addition in A induces an addition

$$(z, w) \longrightarrow z + w = (z_l + w_l, z_r + w_r) \, : \, \mathrm{A}^{(n)} \times \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}$$

for any set in (30). A good Symbol Processor will have designed its built-in addition to carry out such inductive extension in an automatic manner. If the addition endows A with a structure of Abelian group, the addition induced by doubling has the same effect in $\mathrm{A}^{(n)}$ for any n. In which case, if 0 is the neutral element in A, then $0^{(n)}$ is the neutral element in $\mathrm{A}^{(n)}$, i.e., $0^{(n)} + z = z$ for any z in $\mathrm{A}^{(n)}$.

Likewise, a scaling multiplication $(\alpha, x) \longrightarrow \alpha x \, : \, \mathrm{B} \times A \longrightarrow A$ induces by doubling a scaling multiplication

$$(\alpha, z) \longrightarrow \alpha z = (\alpha z_l, \alpha z_r) \, : \, \mathrm{B} \times \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}.$$

One should expect the Symbol Processor to enforce automatically such an induction for its built-in multiplication. In case B is a field and the scaling multiplication endows A with a structure of vector space, the scaling multiplication induced by doubling has the same effect in the sets of (30).

Conjugation brings in a feature that distinguishes the set $\mathrm{A}^{(n)}$ of binary trees of height n from the product set A^{2^n} of 2^n copies of A. This is an operator defined inductively by the rules

$$z \longrightarrow \tilde{z} = \begin{cases} z & \text{for } z \in \mathrm{A}, \\ (\tilde{z}_l, -z_r) & \text{for } z \in \mathrm{A}^{(n)}. \end{cases}$$

The conjugation, one will check easily, is linear; furthermore, it is an involution, that is, if $w = \tilde{z}$ then $\tilde{w} = z$, and that makes it an isomorphism of the vector space $\mathrm{A}^{(n)}$.

With the conjugation one defines 2^n distinct multiplications in $\mathrm{A}^{(n)}$. Let $p = (p_0, p_1, \ldots, p_{n-1})$ be a sequence of scalars all equal either to $+1$ or -1. For each such sequence, consider the binary operator defined inductively by the rules

$$(z, w) \longrightarrow z \otimes_p w = \begin{cases} (z_l w_l + p_0 w_r z_r, z_r w_l + w_r z_l) & \text{for } n = 1, \\ (z_l \otimes_{\check{p}} w_l + p_{n-1} \widetilde{w_r} \otimes_{\check{p}} z_r, \\ \qquad z_r \otimes_{\check{p}} \widetilde{w_l} + w_r \otimes_{\check{p}} z_l) & \text{for } n > 1, \end{cases}$$

where $\check{p}$ is the sequence obtained from p by dropping the last element. The vector space $\mathrm{A}^{(n)}$ with the structure induced by one of the products $\otimes_p$ is called a *hypercomplex system*; p is referred to as the set of *structure constants* for that system.

A few properties of the multiplications of binary trees are obvious: for any $\alpha \in \mathrm{B}$ and for any $z, z', w, w' \in \mathrm{A}^{(n)}$,

$$(\alpha z) \otimes_p w = z \otimes_p (\alpha w) = \alpha (z \otimes_p w);$$

$$z \otimes_p (w + w') = z \otimes_p w + z \otimes_p w',$$

$$(z + z') \otimes_p w = z \otimes_p w + z' \otimes_p w;$$

$$1^{(n)} \otimes_p z = z \otimes_p 1^{(n)} = z.$$

These are precisely the properties a multiplication must enjoy to make an *algebra with a unit element* out of a vector space. When

$$(z \otimes_p z') \otimes_p w = z \otimes_p (z' \otimes_p w) \quad \text{for any } z,\ z' \text{ and } w \ \in \mathrm{A}^{(n)},$$

the algebra is said to be *associative*; when

$$z \otimes_p w = w \otimes_p z \quad \text{for any } z \text{ and } w \ \in \mathrm{A}^{(n)}.$$

it is said to be *commutative*.

Given that the multiplication in A is associative and commutative, the two hypercomplex structures on $\mathrm{A}^{(1)}$ are algebras that are associative as well as commutative. When $\mathrm{A} = \mathbf{R}$, for $p_0 = -1$, $\mathrm{A}^{(1)}$ is the algebra $\mathbf{C}$ of ordinary complex numbers whereas, for $p_0 = 1$, it makes what some authors call the algebra of *hyperbolic* complex numbers. On $\mathrm{A}^{(2)}$, none of the hypercomplex structures is commutative, but all are associative. For $\mathrm{A} = \mathbf{R}$ we recover with $p = (-1, -1)$ the algebra $\mathcal{Q}$ of ordinary quaternions. On $\mathrm{A}^{(3)}$, none of the eight hypercomplex algebras is either associative or commutative, but all of them are *alternative*, that is, for any z and w in $\mathrm{A}^{(3)}$,

$$(z \otimes_p z) \otimes_p w = z \otimes_p (z \otimes_p w),$$

$$z \otimes_p (w \otimes_p w) = (z \otimes_p w) \otimes_p w.$$

The algebra $\mathcal{O}$ of Cayley numbers – also known as *octonions* – corresponds to the case when $\mathrm{A} = \mathbf{R}$ and $p = (-1, -1, -1)$. Beyond $n = 3$, associativity, commutativity and the alternative property are all lost. All the facts we just mentioned are readily checked by Symbol Processors.

For any $n \geq 1$ and any sequence p made of $+1$ and -1, conjugation is a skew involution: for any u and v in $\mathrm{A}^{(n)}$, if $w = u \otimes_p v$, then $\tilde{w} = \tilde{v} \otimes_p \tilde{u}$.

In order to derive a concept of scalar product within the formalism of hypercomplex structures, we need the operator

$$z \longrightarrow \Re(z) \ : \ \mathrm{A}^{(n)} \longrightarrow \mathrm{A}$$

defined inductively by the rules

$$\Re(z) = \begin{cases} z & \text{for } z \in \mathrm{A}, \\ \Re(z_l) & \text{for } z \in \mathrm{A}^{(n)}. \end{cases}$$

Visibly, $\Re$ is a linear form; moreover, for any z in $\mathrm{A}^{(n)}$, $\Re(\tilde{z}) = \Re(z)$, which implies that $2\Re(z) = \Re(z + \tilde{z})$ and justifies calling $\Re(z)$ the *real part* of z. Note also that

$$\Re(z \otimes_p w) = \Re(w \otimes_p z) \quad \text{for any } z \text{ and } w \ \in \mathrm{A}^{(n)},$$

whether or not the multiplication is commutative.

On this form is built the operation

$$(z, w) \longrightarrow z \odot_p w = \Re(z \otimes_p \tilde{w}) \ : \ \mathrm{A}^{(n)} \times \mathrm{A}^{(n)} \longrightarrow A,$$

obviously a symmetric bilinear form, that is to say, for any z and w in $\mathrm{A}^{(n)}$,

$$z \odot_p w = w \odot_p z,$$

$$(\alpha z + \alpha' z') \odot_p w = \alpha(z \odot_p w) + \alpha'(z' \odot_p w).$$

For any $z \in \mathrm{A}^{(n)}$, $z \odot_p 0^{(n)} = 0$; conversely, if $z \odot_p w = 0$ for any $z \in \mathrm{A}^{(n)}$, it is so in particular for each element in the standard base of $\mathrm{A}^{(n)}$, which means that all components of w are 0 or that $w = 0^{(n)}$. Algebraists say in this case that the form $\odot_p$ is *not degenerate*; a symmetric bilinear form that is not degenerate is what they call a *scalar* or *dot* product.

Let it be assumed now that $\mathrm{A} = \mathbf{R}$. A simple reasoning by induction proves that the condition

$$z \odot_p z = 0 \quad \text{for any } z \in \mathbf{R}^{(n)} \Longrightarrow z = 0$$

is satisfied only when all structure constants are equal to -1. Only for those hypercomplex structures over $\mathbf{R}$ does it make sense to introduce the function

$$z \longrightarrow \|z\| = \sqrt{z \odot_p z}.$$

Should we want to make it a *norm* compatible with the structure of algebra induced by $\otimes_p$ on $A^{(n)}$, we must ask that

$$\|z \otimes_p w\| = \|z\| \, \|w\| \quad \text{for any } z \text{ and } w \in \mathbf{R}^{(n)}.$$

According to a famous theorem of Hurwitz (1923), such requirement is satisfied only for $n = 1, 2, 3$, that is to say, in the hypercomplex algebras $\mathbf{C}$ of complex numbers, $\mathcal{Q}$ of quaternions and $\mathcal{O}$ of octonions.

Let it be mentioned that the concept of cross product for quaternions extends at once to any hypercomplex system. It would be defined as the operation

$$(z, w) \longrightarrow z \wedge_p w = \tfrac{1}{2}(z \otimes_p \tilde{w} - w \otimes_p \tilde{z}) \; : \; A^{(n)} \times A^{(n)} \longrightarrow A^{(n)}.$$

Visibly the cross product is a skew bilinear operator. It shares some properties with the cross product of quaternions. With the Symbol Processor, for instance, it is readily checked that

THEOREM 11. – *In the algebras* $A^{(1)}$, $A^{(2)}$ *and* $A^{(3)}$,

$$(u \wedge_p v) \odot_p (v \wedge_p w) = (v \odot_p v)(u \odot_p w) - (u \odot_p v)(v \odot_p w)$$

for any u, v *and* w *and any set* p *of structure constants.*

From the theorem we deduce that the Lagrange identity is verified in the algebras $\mathcal{Q}$ and $\mathcal{O}$:

COROLLARY 11.1. – *In the algebras* $A^{(1)}$, $A^{(2)}$ *and* $A^{(3)}$,

$$(u \wedge_p u) \odot_p (v \wedge_p v) = (u \odot_p u)(v \odot_p v) - (u \odot_p v)^2$$

for any u *and* v *and any set* p *of structure constants.*

It is checked likewise that neither the theorem nor its corollary holds in $A^{(n)}$ for $n \geq 4$.

The doubling process is due to Dickson (1918, p. 158, Eq. (6)) not yet, though, as a general technique for constructing hypercomplex systems and their attendant algebras, solely as a way of passing from complex numbers to Hamilton's quaternions, and from there to Cayley's numbers or octonions. Various authors have taken credit for having given the process its full generality: see, e.g., Kantor and Solodovnikov (1973) and Wene (1984).

Notations create confusions. For instance, the reader should realize that the definition of a hypercomplex product

$$(a_1, a_2)\,(a_3, a_4) = (a_1 a_3 + \mu a_4 \overline{a}_2, \overline{a}_1 a_4 + a_3 a_2)$$

given by Wene (1984, Eq. 2.3) corresponds, in our terms, to the product $(a_3, a_4) \otimes_p (a_1, a_2)$ with factors in inverse order. Our definition of the product agrees with that of Lambert and Kibler (1988).

Hypercomplex systems whose structure constants are not all equal to -1 have their use in theoretical physics: the algebras $\mathbf{R}^{(2)}$ with either $(1,-1)$ or $(-1,1)$ as constant structures figure prominently in the class of Reductions treated by Iwai (1985).

After we completed our coding in *Mathematica,* Anthony Hearn drew our attention to the programming work of Kibler for checking the calculations in the paper he authored with Lambert (1988). Kibler gives no indication whatsoever about his implementation with Hearn's Symbol Processor also known as REDUCE.

3.2. KS-TRANSFORMATIONS BY DOUBLING

We explain here why the KS-transformation can be seen as a doubling of the LC-transformation. Actually we do no more than fitting Vivarelli's treatment to the quaternion version of the KS-transformation. We find here an opportunity for defining her special "anti-involution" in intrinsic terms. Her concept is now open to extensions in higher dimensions through inductive doubling. We pay special attention to the re-labelling of subscripts needed to return from her formulas to Equations (16) in Section 2.

In the context of the preceding subsection, the key feature of our explanation is an operator $z \longrightarrow \hat{z} : \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}$ which we propose to call the *left conjugation.* It is defined by the rules

$$\hat{z} = \begin{cases} z \in \mathrm{A}, & \text{for} \\ (\widetilde{z_l}, z_r) & \text{for } n \in \mathrm{A}^{(n)}. \end{cases}$$

For example,

$$\begin{aligned} \hat{z} &= ((z_0, -z_1), (z_2, z_3)) && \text{for } z \in \mathrm{A}^{(2)}, \\ \hat{z} &= (((z_0, -z_1), (-z_2, -z_3)), ((z_4, z_5), (z_6, z_7))) && \text{for } z \in \mathrm{A}^{(3)}. \end{aligned}$$

Clearly, the left conjugation is a linear operator in $\mathrm{A}^{(n)}$ and an involution and, to boot, a skew involution since, for any set p of structure constants,

$$\widehat{z \otimes_p w} = \hat{w} \otimes_p \hat{z} \quad \text{for any } z \text{ and } w \in \mathrm{A}^{(n)}.$$

Let it be mentioned also that $\hat{\tilde{z}} = \tilde{\hat{z}}$ and that $\Re(\hat{z}) = \Re(z)$. There follows from these facts that the left conjugation preserves the norm, that is to say, that $||\hat{z}|| = ||z||$.

We use the left conjugation to build two kinds of transformations

$$\begin{aligned} z &\longrightarrow \mathcal{H}_p^{(L)}(z) = \hat{z} \otimes_p z \; : \; \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}, \\ z &\longrightarrow \mathcal{H}_p^{(R)}(z) = z \otimes_p \hat{z} \; : \; \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}. \end{aligned}$$

They belong to the classes of what Lambert and Kibler (1988) call respectively *left* and *right Hurwitz* transformations (relative to the left conjugation). For $n = 1$, left and right Hurwitz transformations coincide, on the one hand, with the ordinary LC-transformation

$$(\xi,\eta) \longrightarrow (\xi^2-\eta^2, 2\xi\eta),$$

for $p=(-1)$ and, on the other, with the so-called hyperbolic LC-transformation

$$(\xi,\eta) \longrightarrow (\xi^2+\eta^2, 2\xi\eta)$$

for $p=(1)$. For $n=2$, $\mathcal{H}^{(R)}_{(-1,-1)}$ is determined explicitly by the equations

$$\begin{aligned}
x_0 &= z_0^2+z_1^2-z_2^2-z_3^2,\\
x_1 &= 0,\\
x_2 &= 2(z_0z_2-z_1z_3),\\
x_3 &= 2(z_1z_2+z_0z_3).
\end{aligned}$$

We could accept these equations as equivalent (up to a re-labelling of subscripts in x and z) to the equations (16) set up in Section 2 to define explicitly the KS transformation. Yet, we prefer to take advantage of the resources available in the representation of hypercomplex systems in order to recover exactly the equations (16). We do this most simply by composing the Hurwitz transformations with yet another mapping. First, we define a *reverse* function

$$z \longrightarrow \rho(z) \,:\, \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}$$

by the inductive rules

$$\rho(z) = \begin{cases} z & \text{for } z \in \mathrm{A},\\ (\rho(z_r), \rho(z_l)) & \text{for } z \in \mathrm{A}^{(n)}. \end{cases}$$

Of course, ρ is a linear mapping and an involution, hence an isomorphism of $\mathrm{A}^{(n)}$. Then we build the map

$$z \longrightarrow \psi(z) \,:\, \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}$$

according to the rules

$$\psi(z) = \begin{cases} z & \text{for } n \in \mathrm{A},\\ (\rho(z_l), \widetilde{\rho(z_r)}) & \text{for } n \in \mathrm{A}^{(n)}. \end{cases}$$

Evidently ψ is linear. It is not an involution; nevertheless, it is an isomorphism of $\mathrm{A}^{(n)}$. Indeed, whereas $\psi^m(z)$ is not generally equal to z for $m<4$, $\psi^4(z)=z$ for any $z\in\mathrm{A}^{(n)}$. And now we prescribe the left and right KS-transformations to be the quadratic mappings

$$\begin{aligned}
z &\longrightarrow \mathcal{K}^{(L)}_p(z) = \psi\left(\mathcal{H}^{(L)}_p(z)\right) \,:\, \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)},\\
z &\longrightarrow \mathcal{K}^{(R)}_p(z) = \psi\left(\mathcal{H}^{(R)}_p(z)\right) \,:\, \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}.
\end{aligned}$$

Their explicit equations for $n = 2$ are listed in Table 3.3. Observe there that the equations at the bottom of the second column are identical to the equations (16) of Section 2. We should have anticipated that result had we noticed that, for any $z \in \mathrm{A}^{(2)}$,

$$\psi(z) = e_1 \otimes_{(-1,-1)} \tilde{z},$$

$$\mathcal{K}^{(R)}_{(-1,-1)}(z) = z \otimes_{(-1,-1)} e_1 \otimes_{(-1,-1)} \tilde{z} = z\, e_1\, \overline{z}.$$

There is no room left in this paper for examining what contributions, if any, the seven other KS-transformations make to the problem of linearizing perturbed Keplerian systems. We only wanted here to make the point that the doubling technique we abstracted from Vivarelli's papers makes a handy tool for widening in scope the concept of Kustaanheimo and Stiefel.

3.3. Projective Transformations by Doubling

Vivarelli wanted a doubling of the LC-transformation that yields KS-transformations. We now contribute another way of doubling the LC, this time to produce projective transformations.

Having realized that the ordinary LC-transformation is the mapping

$$z \longrightarrow z \otimes_{(-1)} z \;:\; \mathrm{A}^{(1)} \longrightarrow \mathrm{A}^{(1)}, \tag{31}$$

we extend it for any height n and any set of structure constants into the quadratic transformations

$$z \longrightarrow \mathcal{L}_p(z) = z \otimes_p z \;:\; \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}.$$

Their effects are most concisely described in terms of the linear mapping

$$z \longrightarrow \Im(z) \;:\; \mathrm{A}^{(n)} \longrightarrow \mathrm{A}^{(n)}$$

defined inductively by the rules:

$$\Im(z) = \begin{cases} 0 & \text{for } z \in \mathrm{A}, \\ (\Im(z_l), z_r) & \text{for } z \in \mathrm{A}^{(n)}. \end{cases}$$

In a way, $\Im$ is complementary to $\Re$ in that, for any $z \in \mathrm{A}^{(n)}$,

$$\Re\left(\Im(z)\right) = 0 \quad \text{and} \quad z = \Re(z)\, 1^{(n)} + \Im(z).$$

For these reasons, we refer to $\Im(z)$ as the *imaginary* part of z.

THEOREM 12. – *For any n, any $z \in \mathrm{A}^{(n)}$ and any set p of structure constants,*

$$\Re\left(\mathcal{L}_p(z)\right) = \Re^2(z) - \Im(z) \odot_p \Im(z), \tag{32}$$

$$\Im\left(\mathcal{L}_p(z)\right) = 2\,\Re(z)\,\Im(z), \tag{33}$$

$$\mathcal{L}_p(z) \odot_p \mathcal{L}_p(z) = \left(\Re(z)^2 + \Im(z) \odot_p \Im(z)\right)^2. \tag{34}$$

Left	**Right**
$p = (+1, +1)$	
$x_1 = z_0^2 - z_1^2 + z_2^2 - z_3^2,$	$x_1 = z_0^2 - z_1^2 + z_2^2 - z_3^2,$
$x_2 = -2(z_1 z_2 - z_0 z_3),$	$x_2 = 2(z_1 z_2 + z_0 z_3),$
$x_3 = 2(z_1 z_3 - z_0 z_2)$	$x_3 = -2(z_0 z_2 + z_1 z_3)$
$p = (+1, -1)$	
$x_1 = z_0^2 + z_1^2 + z_2^2 + z_3^2,$	$x_1 = z_0^2 + z_1^2 + z_2^2 + z_3^2,$
$x_2 = -2(z_1 z_2 - z_0 z_3),$	$x_2 = 2(z_1 z_2 + z_0 z_3),$
$x_3 = -2(z_1 z_3 + z_0 z_2)$	$x_3 = 2(z_1 z_3 - z_0 z_2)$
$p = (-1, +1)$	
$x_1 = z_0^2 - z_1^2 - z_2^2 + z_3^2,$	$x_1 = z_0^2 - z_1^2 - z_2^2 - z_3^2,$
$x_2 = -2(z_1 z_2 - z_0 z_3),$	$x_2 = 2(z_1 z_2 + z_0 z_3),$
$x_3 = 2(z_1 z_3 - z_0 z_2)$	$x_3 = -2(z_0 z_2 + z_1 z_3)$
$p = (-1, -1)$	
$x_1 = z_0^2 + z_1^2 - z_2^2 - z_3^2,$	$x_1 = z_0^2 + z_1^2 - z_2^2 - z_3^2,$
$x_2 = -2(z_1 z_2 - z_0 z_3),$	$x_2 = 2(z_1 z_2 + z_0 z_3),$
$x_3 = -2(z_1 z_3 + z_0 z_2)$	$x_3 = 2(z_1 z_3 - z_0 z_2)$

For example, in the algebra $\mathcal{Q}$ of quaternions,

$$\mathcal{L}(z) = (z_0^2 - z_1^2 - z_2^2 - z_3^2, 2z_0 z_1, 2z_0 z_2, 2z_0 z_3)$$

Formulas (32) and (34) can be sharpened when the hypercomplex systems are normed algebras.

THEOREM 13. – *For $n = 2$ and $n = 3$, provided all structure constants are equal to -1, the mapping*

$$(|\Re(z)|, \|\Im(z)\|) \longrightarrow (\Re(\mathcal{L}_p(z)), \|\Im(\mathcal{L}_p(z))\|) \; : \; \mathbf{R}_+ \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+ \times \mathbf{R}$$

is an LC-*transformation.*

Proof. – Indeed, by virtue of (32) and (33),

$$\Re(\mathcal{L}_p(z)) = \Re^2(z) - \|\Im(z)\|^2 \tag{35}$$

and

$$\|\Im(\mathcal{L}_p(z))\| = 2\,|\Re(z)|\,\|\Im(z)\|. \blacksquare$$

From the theorem, we infer that

$$\|\mathcal{L}_p(z)\|^2 = \left(\Re^2(z) - \|\Im(z)\|^2\right)^2 + 4\,\Re^2(z)\,\|\Im(z)\|^2 = \|z\|^4. \tag{36}$$

Furthermore, if ϕ and ψ are the angles such that

$$\begin{aligned} \Re(\mathcal{L}_p(z)) &= \|\mathcal{L}_p(z)\| \cos\phi, & \Re(z) &= \|z\| \cos\psi, \\ \|\Im(\mathcal{L}_p(z))\| &= \|\mathcal{L}_p(z)\| \sin\phi, & \|\Im(z)\| &= \|z\| \sin\psi, \\ 0 \le \phi &\le \pi, & 0 \le \psi &\le \pi/2. \end{aligned} \tag{37}$$

then $\phi = 2\psi$. In other words, as one should expect from an LC-transformation, the mapping mentioned in the theorem "squares the distances and doubles the angles at the origin."

Doubling of a projective decomposition into a Levi-Civita transformation concerns solely the coordinates; the technique offers no hint on how to extend the point-transformation into a canonical transformation. In practice, one tries to find the extension that enhances best some features of the particular problem at hand like its symmetries and integrals, not to mention, of course, the textbook recipe of which we already made an application when, at the suggestion of Kurcheeva, we extended the KS-transformation. Our object now is to use that standard procedure to produce a canonical extension by which the momenta Z and X conjugate respectively to the coordinates z and $x = \mathcal{L}_p(z)$ are required to satisfy identically the differential relation $X \cdot dx = Z \cdot dz$. Provided all structure constants are equal to -1, one will find in this manner that

$$\begin{aligned} &\Re(Z) = 2(z \cdot X), \\ &\Im(Z) = 2\left(\Re(z)\,\Im(X) - \Re(X)\,\Im(z)\right), \end{aligned}$$

and, by inversion, that

$$\begin{aligned}
\Re(X) &= \frac{\Re(z \otimes_p Z)}{\|z\|^2}, \\
\Im(X) &= \frac{1}{2\,\Re(z)} \left[\Im(Z) + \frac{\tilde{z} \odot_p Z}{\|z\|^2} \Im(z) \right] \\
&= \frac{1}{2\,\|z\|^2} \left[\Re(Z)\Im(z) + \frac{1}{\Re(z)} \left(\Im(z) \wedge_p \Im(Z) \right) \wedge_p \Im(z) \right].
\end{aligned} \tag{38}$$

We want to emphasize that these formulas hold not only in four but also in eight dimensions.

4. The projective factoring and its extensions

One may derive satisfaction from the prospect of Stiefel's concept of a KS-transformation and Laplace's technique of projective decomposition stemming from a common root in the theory of hypercomplex systems. In practice, however, the point-transformation in (31) and its canonical extension in (38) do not suit well the common problems in Point Dynamics. One would like, for instance, that $\Re(Z)$ stand for the radial component of the velocity $\Im(X)$ and that $\Im(Z)$ be related in a simple way with the transverse component of $\Im(X)$. Meeting these specifications while maintaining the canonical character is not possible lest a modification in the moment segment of the transformation be compensated by an adjustment in the coordinate segment. Among the specimens of canonical extensions for the projective decomposition in a three-dimensional space, one will thus see posted

$$\Re(x) = \|\Im(z)\|^2$$

for the BF-transformation, and

$$\Re(x) = \Re(z)^2 \left(1 - \|\Im(z)\|^2\right)$$

for the D-transformation instead of (35). More drastic is the stratagem adopted in the DEF-transformation: the canonicity requirements are abandoned in favor of Stiefel's criteria of weak canonicity.

4.1. The DEF-transformation

The DEF-transformation in our terminology is the mapping

$$(u_0, \boldsymbol{u}, U_0, \boldsymbol{U}) \longrightarrow (\boldsymbol{x}, \boldsymbol{X}) : \mathbf{R}_+ \times \mathbf{R}^3 \times \mathbf{R} \times \mathbf{R}^3 \longrightarrow \mathbf{R}^3 \times \mathbf{R}^3$$

defined by the equations

$$\boxed{\begin{aligned}
\boldsymbol{x} &= u_0 \boldsymbol{u}, \\
\boldsymbol{X} &= U_0 \boldsymbol{u} + \frac{1}{u_0} (\boldsymbol{u} \times \boldsymbol{U}) \times \boldsymbol{u}.
\end{aligned}} \tag{39}$$

The classical decomposition $\|\boldsymbol{a}\|^2\,\boldsymbol{b} = (\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{a} + (\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{a}$ of a vector $\boldsymbol{b}$ into its components respectively parallel and perpendicular to a vector $\boldsymbol{a}$ guided our choice for $\boldsymbol{X}$.

As Stiefel and Scheifele did for the KS-transformation, we solve the problem of initial conditions for the DEF-transformation, not a trivial exercise considering that the transformation raises the dimensions and, therefore, is not reversible.

THEOREM 14. – *Given a pair $(\boldsymbol{x}, \boldsymbol{X})$, the elements*

$$u_0 = \frac{\|\boldsymbol{x}\|}{\beta}, \quad U_0 = \frac{\boldsymbol{X}\cdot\boldsymbol{x}}{\beta\|\boldsymbol{x}\|}, \quad \boldsymbol{u} = \frac{\beta}{\|\boldsymbol{x}\|}\boldsymbol{x}, \quad \boldsymbol{U} = \frac{\|\boldsymbol{x}\|}{\beta^3}\boldsymbol{X} - \frac{B}{\|\boldsymbol{x}\|}\boldsymbol{x} \tag{40}$$

dependent on the parameters β and B satisfy the definitions in (39). Conversely, for any quadruple $(u_0, \boldsymbol{u})$ and any quadruple $(U_0, \boldsymbol{U})$, the coordinates and momenta defined in (39) satisfy the relations (40) if and only if

$$\beta = \|\boldsymbol{u}\| \quad \textit{and} \quad B = \frac{1}{\|\boldsymbol{u}\|}(u_0 U_0 - \boldsymbol{u}\cdot\boldsymbol{U}).$$

Proof. – The first part of the theorem is proved simply by replacing the variables u_0, $\boldsymbol{u}$, U_0 and $\boldsymbol{U}$ in the right hand members of (39) by their expressions given in (40). As for the second part, we begin by deducing from (39) that

$$u_0 = \|\boldsymbol{x}\| \,/\, \|\boldsymbol{u}\|,$$

$$\boldsymbol{u} = \frac{1}{u_0}\boldsymbol{x} = \|\boldsymbol{u}\|\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|},$$

$$U_0 = \frac{1}{\|\boldsymbol{u}\|^2}(\boldsymbol{u}\cdot\boldsymbol{X}) = \frac{\boldsymbol{x}\cdot\boldsymbol{X}}{\|\boldsymbol{x}\|\,\|\boldsymbol{u}\|}.$$

Also on account of (39), we find that

$$\boldsymbol{x}\times\boldsymbol{X} = \|\boldsymbol{u}\|^2(\boldsymbol{u}\times\boldsymbol{U}). \tag{41}$$

We regard this relation as an equation in $\boldsymbol{U}$. By virtue of Lemma 5 below, since the vectors $\boldsymbol{u}$ and $\boldsymbol{x}\times\boldsymbol{X}$ are perpendicular, the sum

$$\boldsymbol{U} = \frac{1}{\|\boldsymbol{u}\|^4}(\boldsymbol{x}\times\boldsymbol{X})\times\boldsymbol{u} + \lambda\boldsymbol{u} = \frac{\|\boldsymbol{x}\|}{\|\boldsymbol{u}\|^3}\boldsymbol{X} + \left(\lambda - \frac{u_0 U_0}{\|\boldsymbol{u}\|^2}\right)\boldsymbol{u}$$

represent all possible solutions to equation (41). Taking the dot product of both members by $\boldsymbol{u}$, we find that

$$\lambda = \|\boldsymbol{u}\|^{-2}(\boldsymbol{u}\cdot\boldsymbol{U}).$$

It thus appears that we can adopt

$$\beta = \|\boldsymbol{u}\| \quad \text{and} \quad B = \frac{1}{\|\boldsymbol{u}\|}(u_0 U_0 - \boldsymbol{u}\cdot\boldsymbol{U})$$

as the parameters in (40). ∎

LEMMA 5. – *Let* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *be two linearly independent vectors. The equation* $\boldsymbol{a} \times \boldsymbol{x} = \boldsymbol{b}$ *has a solution if and only if* $\boldsymbol{a} \cdot \boldsymbol{b} = 0$*, in which case the solutions are all linear combinations of the form* $\|\boldsymbol{a}\|^{-2}(\boldsymbol{b} \times \boldsymbol{a}) + \alpha \boldsymbol{a}$ *where* α *is an arbitrary scalar.*

Proof. – By assumption, the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{a} \times \boldsymbol{b}$ constitute a base in the space of three-dimensional real vectors, and any vector $\boldsymbol{x}$ may be decomposed in a unique way into a sum of the form

$$\boldsymbol{x} = \alpha \boldsymbol{a} + \beta \boldsymbol{b} + \gamma(\boldsymbol{a} \times \boldsymbol{b}).$$

There follows that $\boldsymbol{x}$ is a solution of the equation if and only if

$$\beta = 0, \quad \gamma(\boldsymbol{a} \cdot \boldsymbol{b}) = 0 \quad \text{and} \quad 1 + \gamma(\boldsymbol{a} \cdot \boldsymbol{a}) = 0.$$

These conditions are satisfied only when $\boldsymbol{a} \cdot \boldsymbol{b} = 0$. In that case, one must take $\gamma = -\|a\|^{-2}$.∎

In the phase space $(u_0, \boldsymbol{u}, U_0, \boldsymbol{U})$, we adopt the standard symplectic structure so that the Poisson bracket of any pair of smooth functions f and g is the function

$$\{f\, ;\, g\} = \sum_{0 \leq i \leq 3} \left(\frac{\partial f}{\partial u_i} \frac{\partial g}{\partial U_i} - \frac{\partial f}{\partial U_i} \frac{\partial g}{\partial u_i} \right)$$

In contrast to the complexity it encountered with the KS-transformation, the Symbol Processor takes hardly any time for evaluating the Poisson brackets of the DEF-transformation. The results are strikingly simple:

$$\{x_i\, ;\, x_j\} = 0 \quad \text{for} \quad 1 \leq i, j \leq 3,$$

$$\{x_i\, ;\, X_j\} = \delta_{i,j}\, \beta^2 \quad \text{for} \quad 1 \leq i \leq j \leq 3,$$

$$\{X_i\, ;\, X_j\} = 0 \quad \text{for} \quad 1 \leq i, j \leq 3.$$

Let us now show that the DEF-transformation linearizes the equations of motion for Keplerian systems. From the definition (39) we deduce that

$$\|\boldsymbol{X}\|^2 = \beta^2 \left(U_0^2 + \frac{\|\boldsymbol{u} \times \boldsymbol{U}\|^2}{u_0^2} \right).$$

This expression indicates that the angular momenta

$$\boldsymbol{Q} = \boldsymbol{u} \times \boldsymbol{U} \quad \text{and} \quad Q = \|\boldsymbol{Q}\| \tag{42}$$

will figure prominently in the equations of motion. For a clean coding of its partial derivatives with the Symbol Processor, we recommend using the vector identities:

$$\boldsymbol{\nabla}_{\boldsymbol{u}} Q = -(\boldsymbol{Q} \times \boldsymbol{U})/Q \quad \text{and} \quad \boldsymbol{\nabla}_{\boldsymbol{U}} Q = (\boldsymbol{Q} \times \boldsymbol{u})/Q.$$

With these notations, applying the DEF-transformation to the Keplerian (26) will produce the function

$$\mathcal{K} = \frac{\beta^2}{2}\left(U_0^2 + \frac{Q^2}{u_0^2}\right) - \frac{\mu}{\beta u_0}.$$

The "new" equations of motion are then

$$\begin{aligned}
\frac{du_0}{dt} &= \{u_0\,;\,\mathcal{K}\} = \beta^2 U_0,\\
\frac{dU_0}{dt} &= \{U_0\,;\,\mathcal{K}\} = \beta^2\left(\frac{Q^2}{u_0^3} - \frac{\mu}{\beta^3 u_0^2}\right),\\
\frac{d\boldsymbol{u}}{dt} &= \{\boldsymbol{u}\,;\,\mathcal{K}\} = \frac{\beta^2}{u_0^2}\boldsymbol{Q}\times\boldsymbol{u},\\
\frac{d\boldsymbol{U}}{dt} &= \{\boldsymbol{U}\,;\,\mathcal{K}\} = \frac{\beta^2}{u_0^2}\boldsymbol{Q}\times\boldsymbol{U} - \frac{1}{\beta^2}\left(2\mathcal{H} + 3\frac{\mu}{\beta u_0}\right)\boldsymbol{u}.
\end{aligned} \tag{43}$$

As is done usually for Keplerian systems, we replace the independent variable t by a *generalized* true anomaly f such that

$$u_0^2\,df = \beta^2 Q\,dt.$$

We deduce readily from the first equation in (43) that

$$\frac{d^2\boldsymbol{u}}{df^2} = \frac{d}{df}\left(\frac{\boldsymbol{Q}\times\boldsymbol{u}}{Q}\right) = \frac{\boldsymbol{Q}}{Q}\times\frac{d\boldsymbol{u}}{df} = \frac{\boldsymbol{Q}}{Q^2}\times(\boldsymbol{Q}\times\boldsymbol{u}) = -\boldsymbol{u}$$

while, for the (dimensionless) variable

$$\sigma = Q^2/(\mu u_0), \tag{44}$$

we find that

$$\begin{aligned}
\frac{d\sigma}{df} &= -\frac{Q}{\mu}U_0,\\
\frac{d^2\sigma}{df^2} &= -\frac{Q}{\mu}\frac{dU_0}{df} = -\sigma + \frac{1}{\beta^3}.
\end{aligned} \tag{45}$$

There remains to establish the correspondence between the solutions of the canonical equations derived from (26) and those of the system (43). The presence of the factor β^2 in the Poisson brackets $\{x_i\,;\,X_j\}$ calls for restricting the vector $\boldsymbol{u}$ to the sphere $\beta = \|\boldsymbol{u}\| = 1$, which condition establishes the DEF-transformation as a weakly canonical extension on the manifold $\mathcal{S} = \mathbf{R}_+ \times \mathbf{S}^2 \times \mathbf{R} \times \mathbf{R}^3$ in the space $(u_0, \boldsymbol{u}, U_0, \boldsymbol{U})$. On the hypersurface $\mathcal{S}$, $\boldsymbol{u}$ and u_0 are the radial direction and the radial distance respectively while U_0 is the radial component of the momentum vector $\boldsymbol{X}$. The restriction to $\mathcal{S}$ is acceptable in Dynamics because the Poisson brackets

$$\{\beta\,;\,x_i\} = \{\beta\,;\,X_i\} = 0 \quad \text{for} \quad 1 \leq i \leq 3,$$

imply that $\{\beta\,;\,F\} = 0$ for any smooth function $F(\boldsymbol{x}, \boldsymbol{X})$, and this, in turn, means that β is an integral for any Hamiltonian system in the phase space $(\boldsymbol{x}, \boldsymbol{X})$. We must also add that the restriction to $\mathcal{S}$ has the good effect of changing the second equation in (45) into a linear differential equation

$$\left(\frac{d^2}{df^2} + 1\right)\sigma = 1$$

with constant coefficients. This completes the linearization of Keplerian systems in three dimensions by means of the DEF-transformation.

4.2. Canonical Embeddings

Lowering the dimensions of a Hamiltonian system is a step most frequent in Mechanics. Given a Hamiltonian $\mathcal{H}(\boldsymbol{p}, \boldsymbol{P})$ in n coordinates $\boldsymbol{p}$ and their conjugate momenta $\boldsymbol{P}$, one creates a canonical transformation

$$\gamma : (\boldsymbol{q}, \boldsymbol{Q}) \longrightarrow (\boldsymbol{p}, \boldsymbol{P}) : \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R}^n \times \mathbf{R}^n$$

that will render m coordinates, say $(q_i)_{1 \leq i \leq m}$, ignorable in the pullback of $\mathcal{H}$, that is, the function

$$\gamma^*\mathcal{H}(\boldsymbol{q}, \boldsymbol{Q}) = \mathcal{H}(\boldsymbol{p}(\boldsymbol{q}, \boldsymbol{Q}), \boldsymbol{P}(\boldsymbol{q}, \boldsymbol{Q})).$$

The transformation γ is designed to reveal that the conjugate momenta $(Q_i)_{1 \leq i \leq m}$ are integrals of the system; the general technique known as Reduction uses them to lower the dimension of the original system by $2m$. Not until the publication of *Linear and Regular Celestial Mechanics* had the reverse problem, that of increasing the dimensions of a Hamiltonian system, been explored somehow in a systematic manner.

The framework set up by Stiefel and Scheifele is well known (1971, Theorem 2, pp. 189–190). Assume that γ maps a domain of dimension $2(n+m)$ in the phase space $(\boldsymbol{q}, \boldsymbol{Q})$ into a domain of dimension $2n$ in the phase space $(\boldsymbol{p}, \boldsymbol{P})$, and define the Poisson bracket $\{f\,;\,g\}$ of any pair of smooth functions f and g of $(\boldsymbol{p}, \boldsymbol{P})$ as the function

$$\{f\,;\,g\} = \nabla_{\boldsymbol{q}}(\gamma^* f) \cdot \nabla_{\boldsymbol{Q}}(\gamma^* g) - \nabla_{\boldsymbol{Q}}(\gamma^* f) \cdot \nabla_{\boldsymbol{q}}(\gamma^* g).$$

Provided that

$$\{p_i\,;\,p_j\} = \{P_i\,;\,P_j\} = 0 \quad \text{for} \quad 1 \leq i \leq j \leq n, \tag{46}$$

and that there exists a so-called multiplier μ independent of $\boldsymbol{q}$ and $\boldsymbol{Q}$ such that

$$\{p_i\,;\,P_j\} = \mu\delta_{i,j} \quad \text{for} \quad 1 \leq i \leq j \leq n \tag{47}$$

where $\delta_{i,j}$ is the Kronecker symbol, then

$$\gamma^*\boldsymbol{p}(t) = \boldsymbol{p}(\boldsymbol{q}(t), \boldsymbol{Q}(t)) \quad \text{and} \quad \gamma^*\boldsymbol{P}(t) = \boldsymbol{P}(\boldsymbol{q}(t), \boldsymbol{Q}(t))$$

is a solution of the canonical system

$$\dot{\boldsymbol{p}} = \boldsymbol{\nabla}_{\boldsymbol{P}}\mathcal{H}, \quad \dot{\boldsymbol{P}} = -\boldsymbol{\nabla}_{\boldsymbol{p}}\mathcal{H}$$

whenever $(\boldsymbol{q}(t), \boldsymbol{Q}(t))$ is a solution of the system

$$\dot{\boldsymbol{q}} = \boldsymbol{\nabla}_{\boldsymbol{Q}}\Big(\gamma^*(\mu^{-1}\mathcal{H})\Big), \quad \dot{\boldsymbol{Q}} = -\boldsymbol{\nabla}_{\boldsymbol{q}}\Big(\gamma^*(\mu^{-1}\mathcal{H})\Big).$$

Of a transformation γ with properties (46) and (47) we say that it is *weakly* canonical. For a slight generalization of the condition (47), the reader is referred to the Appendix.

The theorem of Stiefel and Scheifele calls for complements in two directions. Firstly, it does not supply a technique for building canonical transformations capable of raising the dimension of a Hamiltonian system – something that Liouville and Routh do not do either for lowering the dimension. Usually, though, one starts with a coordinate transformation

$$\gamma^\flat : \boldsymbol{q} \longrightarrow \boldsymbol{p} = \boldsymbol{p}(\boldsymbol{q})$$

of a domain in a coordinate space $(\boldsymbol{q})$ of dimension $n + m$ into a domain in a coordinate space $(\boldsymbol{p})$ of dimension n, in which case Kurcheeva (1977) has provided an algorithm for building some types of canonical extensions of $\gamma^\flat$ which satisfy the conditions of the theorem. We have applied Kurcheeva's technique in Sections 2 and 3. Secondly, Stiefel and Scheifele are silent on the question of establishing the complementary between the raising and the lowering of dimensions for Hamiltonian systems. Lidov (1982) was the first to broach the issue. Evidently, if $\gamma^*\mathcal{H}$ is the Hamiltonian obtained from $\mathcal{H}$ by reduction through a canonical transformation in the classical sense, then one can look back at $\mathcal{H}$ as having been obtained from $\gamma^*\mathcal{H}$ by increasing the dimension of the system. But what is of interest here is the reverse question: given $\mathcal{H}$ in a phase space $(\boldsymbol{p}, \boldsymbol{P})$ of dimension $2n$, construct a transformation γ that is canonical in the classical sense, raises the dimension of the system by $2m$ units, preserves its canonical character, and gets the "old" Hamiltonian to result from its pullback $\gamma^*\mathcal{H}$ by reduction. Practically, starting from a transformation

$$\boldsymbol{q} \longrightarrow \boldsymbol{p} = \boldsymbol{p}(\boldsymbol{q})$$

with $\boldsymbol{p} \in \mathbf{R}^n$ and $\boldsymbol{q} \in \mathbf{R}^{n+m}$ that raises the dimension of the coordinate space, one must find a vector function $\boldsymbol{P}(\boldsymbol{q}, \boldsymbol{Q})$ of dimension n and also two vector functions $\boldsymbol{r}(\boldsymbol{q}, \boldsymbol{Q})$ and $\boldsymbol{R}(\boldsymbol{q}, \boldsymbol{Q})$ of dimension m to define a canonical homeomorphism

$$(\boldsymbol{q}, \boldsymbol{Q}) \longrightarrow (\boldsymbol{p}, \boldsymbol{r}, \boldsymbol{P}, \boldsymbol{R})$$

that satisfies the requirements. If this is feasible, then one can say that the dimension increasing transformation has rendered the coordinates $\boldsymbol{r}$ ignorable, or that the old Hamiltonian $\mathcal{H}$ is obtained from $\gamma^*\mathcal{H}$ by reduction through the integrals $\boldsymbol{R}$.

4.3. The D-transformation

As a case in point, let us consider the transformation

$$(u_0, \boldsymbol{u}, U_0, \boldsymbol{U}) \longrightarrow (\boldsymbol{x}, \boldsymbol{X}) : \mathbf{R}_+ \times \mathbf{R}^3 \times \mathbf{R} \times \mathbf{R}^3 \longrightarrow \mathbf{R}^3 \times \mathbf{R}^3$$

defined by the equations

$$\boxed{\begin{aligned} \boldsymbol{x} &= u_0 \boldsymbol{u}, \\ \boldsymbol{X} &= U_0 \boldsymbol{u} + \frac{1}{u_0}\left[\boldsymbol{U} - (\boldsymbol{u}.\boldsymbol{U})\boldsymbol{u}\right]. \end{aligned}} \tag{48}$$

We call it the D-transformation. With a Symbol Processor one checks easily that, for $1 \leq i \leq j \leq 3$,

$$\{x_i \,;\, x_j\} = \{X_i \,;\, X_j\} = 0 \quad \text{and} \quad \{x_i \,;\, X_j\} = \delta_{i,j},$$

that is to say, that the D-transformation is weakly canonical. To make it properly canonical as a mapping of an eight-dimensional space into an eight-dimensional space, a coordinate and a momentum must be added to $\boldsymbol{x}$ and $\boldsymbol{X}$. Finding these new variables is a matter of trials and errors with the Symbol Processor. Grouping the many terms arising from a Poisson bracket into combinations easily identifiable is the most delicate part of the assignment. For instance, given an arbitrary smooth function F of $\boldsymbol{x}$ and $\boldsymbol{X}$, we found that

$$\begin{aligned} \{u_0 \,;\, F\} &= (\boldsymbol{u} \cdot \boldsymbol{\nabla}_{\boldsymbol{X}} F), \\ \left\{\tfrac{1}{2}(\boldsymbol{u} \cdot \boldsymbol{u}) \,;\, F\right\} &= u_0^{-1}(1 - \|\boldsymbol{u}\|^2)(\boldsymbol{u} \cdot \boldsymbol{\nabla}_{\boldsymbol{X}} F). \end{aligned}$$

From these identities we inferred that, for any constant a,

$$\{(1 - \|\boldsymbol{u}\|^2)u_0^a \,;\, F\} = (2 - a)u_0^{a-1}(1 - \|\boldsymbol{u}\|^2)(\boldsymbol{u} \cdot \boldsymbol{U})\,\boldsymbol{\nabla}_{\boldsymbol{X}} F.$$

In particular, for $a = 2$, there follows that $\{u_0^2(1 - \boldsymbol{u} \cdot \boldsymbol{u}) \,;\, F\} = 0$. Likewise, combining the formulas

$$\begin{aligned} \{u_0 U_0 \,;\, F\} &= \boldsymbol{X} \cdot \boldsymbol{\nabla}_{\boldsymbol{X}} F - \boldsymbol{x} \cdot \boldsymbol{\nabla}_{\boldsymbol{x}} F, \\ \{\boldsymbol{u} \cdot \boldsymbol{U} \,;\, F\} &= (2u_0^{-1}\boldsymbol{U} - \boldsymbol{X}) \cdot \boldsymbol{\nabla}_X F - \boldsymbol{x} \cdot \boldsymbol{\nabla}_{\boldsymbol{x}} F, \end{aligned}$$

we came to realize that

$$\{u_0^a(u_0 U_0 - \boldsymbol{u} \cdot \boldsymbol{U}) \,;\, F\} = (2 + a)u_0^{a-1}(u_0 U_0 - \boldsymbol{u} \cdot \boldsymbol{U})(\boldsymbol{u} \cdot \boldsymbol{\nabla}_X F).$$

Therefore, taking $a = -2$, we concluded that

$$\left\{u_0^{-2}(u_0U_0 - \boldsymbol{u}\cdot\boldsymbol{U})\,;\; F\right\} = 0.$$

A simple code in Mathematica then enabled us to check that

$$\left\{au_0^2(1 - \|\boldsymbol{u}\|^2)\,;\; bu_0^{-2}(u_0U_0 - \boldsymbol{u}\cdot\boldsymbol{U})\right\} = 2ab.$$

In particular, for $a = 1$ and $b = 1/2$, the functions

$$\begin{aligned} x_0 &= x_0(u_0, \boldsymbol{u}) = u_0^2(1 - \|\boldsymbol{u}\|^2),\\ X_0 &= X_0(u_0, \boldsymbol{u}, U_0, \boldsymbol{U}) = \tfrac{1}{2}u_0^{-2}(u_0U_0 - \boldsymbol{u}\cdot\boldsymbol{U}) \end{aligned} \tag{49}$$

are such that $\{x_0\,;\,X_0\} = 1$. This statement concludes the proof that the transformation

$$(u_0, \boldsymbol{u}, U_0, \boldsymbol{U}) \longrightarrow (\boldsymbol{x}, x_0, \boldsymbol{X}, X_0) : \mathbf{R}_+ \times \mathbf{R}^3 \times \mathbf{R} \times \mathbf{R}^3 \longrightarrow \mathbf{R}^4 \times \mathbf{R}^4$$

is canonical in the full sense of the word. On the one hand, the transformation raises the dimension of any dynamical system defined over the phase space $(\boldsymbol{x}, \boldsymbol{X})$; on the other hand, it sets x_0 and X_0 as integrals for the pullback Hamiltonian $\gamma^*\mathcal{H}$ over the phase space $(u_0, \boldsymbol{u}, U_0, \boldsymbol{U})$, thereby giving to think that the 'old' system proceeds from the 'new' Hamiltonian $\gamma^*\mathcal{H}$ by reduction.

The D-transformation being invertible, we can look for its inverse. Solving the system made of equations (48) and (49) makes no complications save for the momentum $\boldsymbol{U}$. We handle that case, as we did when we solved the initial value problem for the DEF-transformation, by deducing from (48) that

$$\boldsymbol{x} \times \boldsymbol{X} = \boldsymbol{u} \times \boldsymbol{U},$$

then applying Lemma 5 to solve this equation. The equations of the inverse transformation, namely

$$\begin{aligned} u_0 &= \sqrt{\|\boldsymbol{x}\|^2 + x_0}, & \boldsymbol{u} &= \frac{\boldsymbol{x}}{\sqrt{\|\boldsymbol{x}\|^2 + x_0}},\\ U_0 &= \frac{\boldsymbol{x}\cdot\boldsymbol{X} + 2x_0X_0}{\sqrt{\|\boldsymbol{x}\|^2 + x_0}}, & \boldsymbol{U} &= (\boldsymbol{X} - 2X_0\boldsymbol{x})\sqrt{\|\boldsymbol{x}\|^2 + x_0}, \end{aligned} \tag{50}$$

say that, only at the intersection of the manifolds $x_0 = 0$ and $X_0 = 0$, does u_0 represent the central distance and $\boldsymbol{u}$ the radial direction; furthermore, if $\boldsymbol{X}$ stands for the velocity, then U_0 stands there for the radial velocity.

To our surprise we found that the D-transformation is nothing but the canonical extension of the point-transformation $(u_0, \boldsymbol{u}) \longrightarrow (x_0, \boldsymbol{x})$ defined by the first equation in (48) and the first equation in (49). Indeed it satisfies the differential identity

$$X_0\, dx_0 + \boldsymbol{X} \cdot d\boldsymbol{x} = U_0\, du_0 + \boldsymbol{U} \cdot d\boldsymbol{u}.$$

Furthermore, the momenta

$$\begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial x_0}{\partial u_0} & \dfrac{\partial x_1}{\partial u_0} & \dfrac{\partial x_2}{\partial u_0} & \dfrac{\partial x_3}{\partial u_0} \\ \dfrac{\partial x_0}{\partial u_1} & \dfrac{\partial x_1}{\partial u_1} & \dfrac{\partial x_2}{\partial u_1} & \dfrac{\partial x_3}{\partial u_1} \\ \dfrac{\partial x_0}{\partial u_2} & \dfrac{\partial x_1}{\partial u_2} & \dfrac{\partial x_2}{\partial u_2} & \dfrac{\partial x_3}{\partial u_2} \\ \dfrac{\partial x_0}{\partial u_3} & \dfrac{\partial x_1}{\partial u_3} & \dfrac{\partial x_2}{\partial u_3} & \dfrac{\partial x_3}{\partial u_3} \end{pmatrix} \begin{pmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

obtained from this differential relation are identical to those obtained in (50) by inversion of the D-transformation.

There remains to show that the D-transformation linearizes the Keplerian systems. To this end, we apply it to the Hamiltonian (26). A few straightforward manipulations will decompose the new Hamiltonian into the sum

$$\begin{aligned} \mathcal{K} = (\boldsymbol{u} \cdot \boldsymbol{u}) \left[\frac{1}{2} \left(U_0^2 + \frac{Q^2}{u_0^2} \right) + \frac{x_0}{u_0^4} (\boldsymbol{U} \cdot \boldsymbol{U}) \right] - \frac{\mu}{r} \\ + \frac{x_0}{u_0^2} \left[\frac{x_0}{2u_0^4} (\boldsymbol{U} \cdot \boldsymbol{U}) + 2\, X_0 (\boldsymbol{u} \cdot \boldsymbol{U}) \right]. \end{aligned}$$

The equations of motion in the phase variables $(u_0, \boldsymbol{u}, U_0, \boldsymbol{U})$ are now:

$$\begin{aligned} \frac{du_0}{dt} &= \{u_0\,;\, \mathcal{K}\} = \boldsymbol{u} \cdot \boldsymbol{X}, \\ \frac{dU_0}{dt} &= \{U_0\,;\, \mathcal{K}\} = \frac{1}{u_0} \left(\frac{Q^2}{u_0^2} - \frac{\mu}{r} \right) + \frac{x_0}{u_0^3} \left[2X_0 (\boldsymbol{u} \cdot \boldsymbol{U}) + \frac{\boldsymbol{U} \cdot \boldsymbol{U}}{u_0^2} \right], \\ \frac{d\boldsymbol{u}}{dt} &= \{\boldsymbol{u}\,;\, \mathcal{K}\} = \frac{\boldsymbol{Q} \times \boldsymbol{u}}{u_0^2} + \frac{x_0}{u_0^3} \boldsymbol{X}, \\ \frac{d\boldsymbol{U}}{dt} &= \{\boldsymbol{U}\,;\, \mathcal{K}\} = \frac{\boldsymbol{Q} \times \boldsymbol{U}}{u_0^2} + \left(\frac{\boldsymbol{U} \cdot \boldsymbol{U}}{u_0^2} - \frac{\mu u_0^2}{r^3} \right) \boldsymbol{u} \\ &\qquad - 2X_0 \left(\frac{x_0}{u_0^2} \boldsymbol{U} + 2X_0 u_0^2 \boldsymbol{u} \right). \end{aligned} \tag{51}$$

To the D-transformation we associate the *generalized* true anomaly f defined by the differential form

$$u_0^2\, df = Q\, dt.$$

Accordingly, for σ as defined in (44) and the direction $\boldsymbol{u}$, we deduce from the above system that

$$\begin{aligned} \left(\frac{d^2}{df^2}+1\right)\sigma &= \frac{1}{\|\boldsymbol{u}\|} - \frac{x_0}{\mu u_0}(\boldsymbol{X}\cdot\boldsymbol{X}),\\ \left(\frac{d^2}{df^2}+1\right)\boldsymbol{u} &= -\frac{x_0}{Q^2}\left(\boldsymbol{X}\cdot\boldsymbol{X}+\frac{\mu u_0^2}{r^3}\right)\boldsymbol{u}. \end{aligned} \tag{52}$$

The function x_0 being an integral for any Hamiltonian system represented by a function dependent solely on $\boldsymbol{x}$ and $\boldsymbol{X}$, it is allowed to restrict the D-transformation to the manifold $x_0 = 0$. On that hypersurface, $\|\boldsymbol{u}\| = 1$, and the equations in (52) turn out to be the linear system with constant coefficients:

$$\left(\frac{d^2}{df^2}+1\right)\sigma = 1, \quad \left(\frac{d^2}{df^2}+1\right)\boldsymbol{u} = 0$$

familiar – we recalled in Section 1 – to Laplace and his predecessors.

4.4. The Burdet-Ferrándiz transformation

Ferrándiz gave two versions of the transformation:

$$\boldsymbol{x} = \frac{\boldsymbol{u}}{\rho}, \qquad \boldsymbol{X} = \frac{\rho}{\|\boldsymbol{u}\|^2}\left[(\boldsymbol{u}\times\boldsymbol{U})\times\boldsymbol{u} - \rho P\boldsymbol{u}\right],$$

and

$$\boxed{\begin{aligned} \boldsymbol{x} &= u_0\boldsymbol{u},\\ \boldsymbol{X} &= \frac{1}{\|\boldsymbol{u}\|^2}\left[U_0\boldsymbol{u} + \frac{1}{u_0}(\boldsymbol{u}\times\boldsymbol{U})\times\boldsymbol{u}\right]. \end{aligned}} \tag{53}$$

We refer to the latter as the BF-transformation. One goes from one version to the other through the canonical transformation $(\rho, P) \longrightarrow (u_0, U_0)$ defined by the equations

$$\rho\, u_0 = 1 \quad \text{and} \quad P\rho + U_0 u_0 = 0.$$

Whereas Burdet and Ferrándiz worked principally with the first version – the original one –, we favor the second one for it enhances the affinity of the BF-transformation with the doubling $\mathcal{L}_p$ manufactured in Section 3.3 as well as with the DEF- and D-transformations of the preceding subsections.

With a Symbol Processor, one checks readily two statements proved by Ferrándiz (1988, Lemma 1 p. 347 and Theorem 2 p. 348), namely,

(i) that the BF-transformation is weakly canonical, which statement we recall means that, for $1 \leq i \leq j \leq 3$,

$$\{\boldsymbol{x}_i\,;\,\boldsymbol{x}_j\} = \{\boldsymbol{X}_i\,;\,\boldsymbol{X}_j\} = 0 \quad \text{and} \quad \{\boldsymbol{x}_i\,;\,\boldsymbol{X}_j\} = \delta_{i,j};$$

(ii) that, given the functions

$$x_0 = (\boldsymbol{u}\cdot\boldsymbol{u}) \quad \text{and} \quad B = \boldsymbol{u}\cdot\boldsymbol{U} - u_0 U_0,$$

the relations

$$\{x_0\,;\,F(\boldsymbol{x},\boldsymbol{X})\} = \{B, F(\boldsymbol{x}\,;\,\boldsymbol{X})\} = 0$$

hold uniformly in $(\boldsymbol{u},\boldsymbol{U})$ for any smooth function F by virtue of the BF-transformation.

We want to sharpen the way Ferrándiz characterizes the BF-transformation (i) by completing it to make of it a standard canonical map over an eight-dimensional phase space, and (ii) by verifying that the moment segment in the completed map is the Mathieu extension of its coordinate segment.

To this end, having observed that $\{x_0\,;\,B\} = 2(\boldsymbol{u}\cdot\boldsymbol{u})$, we introduced the function

$$X_0 = \frac{B}{2x_0} = \frac{1}{2x_0}(\boldsymbol{u}\cdot\boldsymbol{U} - u_0 U_0)$$

so as to have that $(x_0, X_0) = 1$. As it was the case for the D-transformation, we are now in a position to embed the BF-transformation into a completely canonical mapping

$$(u_0, \boldsymbol{u}, U_0, \boldsymbol{U}) \longrightarrow (x_0, \boldsymbol{x}, X_0, \boldsymbol{X}).$$

We found that the inverse of that canonical mapping is defined by the equations

$$u_0 = \frac{\|\boldsymbol{x}\|}{\sqrt{x_0}}, \qquad \boldsymbol{u} = \boldsymbol{x}\frac{\sqrt{x_0}}{\|\boldsymbol{x}\|},$$

$$U_0 = (\boldsymbol{x}\cdot\boldsymbol{X})\frac{\sqrt{x_0}}{\|\boldsymbol{x}\|}, \qquad \boldsymbol{U} = \boldsymbol{X}\frac{\|\boldsymbol{x}\|}{\sqrt{x_0}} + 2X_0\boldsymbol{x}\frac{\sqrt{x_0}}{\|\boldsymbol{x}\|}.$$

Then we checked that these are precisely the equations we would obtain by choosing the momenta U_0 and $\boldsymbol{U}$ as we did for the D-transformation in Section 4.2 to satisfy the differential form

$$X_0\,dx_0 + \boldsymbol{X}\cdot d\boldsymbol{x} = U_0\,du_0 + \boldsymbol{U}\cdot d\boldsymbol{u},$$

which goes to show that the duly completed BF-transformation is the canonical extension of a point-transformation.

Applying the BF-transformation to the Hamiltonian (26) yields the function

$$\mathcal{K} = \frac{1}{2x_0}\left(U_0^2 + \frac{Q^2}{u_0^2}\right) - \frac{\mu}{r}$$

and the equations of motion in the new variables:

$$\begin{aligned} \dot{u}_0 &= \frac{1}{x_0}U_0, & \dot{\boldsymbol{u}} &= \frac{1}{x_0 u_0^2}\boldsymbol{Q}\times\boldsymbol{u}, \\ \dot{U}_0 &= \frac{1}{x_0}\left(\frac{Q^2}{u_0^3} - \frac{\mu\sqrt{x_0}}{u_0^2}\right), & \dot{\boldsymbol{U}} &= \frac{1}{x_0 u_0^2}\boldsymbol{Q}\times\boldsymbol{U} + \frac{\boldsymbol{u}}{x_0}\left(2\mathcal{K} + \frac{\mu}{r}\right). \end{aligned} \tag{54}$$

As the *generalized* true anomaly associated with the BF-transformation, Ferrándiz (1988, Eq. 34 p. 353) chooses an angle f such that

$$x_0 u_0^2\, df = Q\, dt.$$

It is somewhat easy with a Symbol Processor to establish that the equations (54) are equivalent to the system

$$\left(\frac{d^2}{df^2} + 1\right)\sigma = \sqrt{x_0}, \qquad \left(\frac{d^2}{df^2} + 1\right)\boldsymbol{u} = 0$$

where σ is the variable defined in (44). We conclude this review of the Burdet-Ferrándiz transformation by noting that, on the integral manifold $x_0 = 1$, the latter system turns to be identical to the classical linear equations for a Keplerian system.

Conclusions

Three themes are weaved on the weft of these Notes. The first concerns the construction of canonical extensions for dimension-raising point-transformations, which, as one should have expected, remains an exercise of improvisation in craftsmanship. The second theme turns on the eighteenth century practice of decomposing the motion of a mass point into a radial displacement and a rotation of the radial direction. For good or ill, that heritage fell into oblivion. The account given in the Notes emphasizes the modernity of prevectorial Dynamics without turning a blind eye to the genuinely cumbersome and archaic. Finally these Notes attempt to confront directly the painful problems of algebraic complexity. Anxious lest they find themselves caught in insufferably long plots of simplification machinations and other algebraic intrigues, experts in celestial mechanics have erred on the side of squeamishness in dealing with this issue. We have returned it to the center of the story since it seems to us that it is not merely the disagreeable tool by which other more substantial ends are accomplished. In some exhilaratingly unavoidable sense, extensive algebraic calculations belong to the processes of creativity in celestial mechanics. As Symbol Processors grow smarter from year to year, the more insistently they will beckon their devotees toward projects deemed "not very promising" by the standards of today.

Acknowledgements

Support for this research came in part from the Applied and Computational Mathematics Program of the Defense Advanced Research Project Agency (Washington), from the Spanish Ministerio de Educación y Ciencia (grant CICYT ESP91-0919 and program Propio), and from the Département de mathématiques spatiales at the Centre National d'Etudes Spatiales (Toulouse). Thanks are due to Professor Ferrándiz for informations about the origin of his transformation, and to Lieutenant Teodoro López of the Real Instituto y Observatorio de la Armada (San Fernando) for a thorough check of the paper in its last draft.

Authors are listed in alphabetical order.

Appendix

Stiefel and Scheifele proved more than what they stated (1971, §31) about dimension raising transformations. With just one or two minor modifications in the demonstration of Stiefel and Scheifele, one easily proves the following

THEOREM 15. – *Given the transformation*

$$(\boldsymbol{q}, \boldsymbol{Q}) \longrightarrow (\boldsymbol{p}, \boldsymbol{P}) : \mathbf{R}^{m+n} \times \mathbf{R}^{m+n} \longrightarrow \mathbf{R}^n \times \mathbf{R}^n,$$

assume that the corresponding Jacobian matrix has the maximum rank $2n$ *and that the Poisson brackets with respect to* $\boldsymbol{q}$ *and* $\boldsymbol{Q}$ *satisfy the relations*

$$\{p_i\,;\,p_j\} = 0, \quad \{p_i\,;\,P_j\} = \Phi\,\delta_{i,j}, \quad \{P_i\,;\,P_j\} = 0 \quad \textit{for} \quad 1 \le i \le j \le n,$$

where Φ *is a function of* $\boldsymbol{q}$ *and* $\boldsymbol{Q}$. *For any Hamiltonian* $\mathcal{H}(\boldsymbol{p}, \boldsymbol{P})$, *let* $\mathcal{K}$ *be the function obtained by inserting the transformation into* $\mathcal{H}$. *If the functions* $\boldsymbol{q}(s)$ *and* $\boldsymbol{Q}(s)$ *satisfy the differential system*

$$\frac{d\boldsymbol{q}}{ds} = \frac{\partial \mathcal{K}}{\partial \boldsymbol{Q}}, \qquad \frac{d\boldsymbol{Q}}{ds} = -\frac{\partial \mathcal{K}}{\partial \boldsymbol{q}},$$

then the functions $\boldsymbol{p}(t)$ *and* $\boldsymbol{P}(t)$ *obtained from* $\boldsymbol{q}(s)$ *and* $\boldsymbol{Q}(s)$ *by the transformation satisfy the canonical equations*

$$\frac{d\boldsymbol{p}}{dt} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{P}}, \qquad \frac{d\boldsymbol{P}}{dt} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}},$$

provided the independent variables s *and* t *are linked by the differential identity*

$$dt = \Phi\, ds.$$

References

Abad, A., Palacios, M. and Sein-Echaluce, M.: 1991, "An efficient formulation for numerical orbit computation," in *Proceedings of the 3rd International Symposium on Spacecraft Flight Dynamics*, ESOC Darmstadt, 65–72.

Bohlin, K.: 1911, "Note sur le problème des deux corps et sur une intégration nouvelle dans le problème des trois corps," *Bull. Astron.* **28**, 113–119.
Bond, V. R.: 1985, "A transformation of the two-body problem," *Celes. Mech.* **35**, 1–7.
Brown, E. W.: 1896, *An Introductory Treatise on the Lunar Theory*. Cambridge, The University Press.
Brown, E. W. and Shook, C. A.: 1933, *Planetary Theory*. Cambridge, The University Press.
Burdet, C. A.: 1967, "Regularization of the two-body problem," *Z. angew. Math. Phys.* **18**, 434–438.
Burdet, C. A.: 1968, "Theory of Kepler motion: The general perturbed two-body problem," *Z. angew. Math. Phys.* **19**, 345–368.
Burdet, C. A.: 1969, "Le mouvement képlérien et les oscillateurs harmoniques," *J. Reine Angew. Math.* **238**, 71–84.
Darboux, G.: 1915, *Leçons sur la théorie générale des surfaces*, Tome I. Paris, Gauthier-Villars.
Deprit, A. and Williams, C. A.: 1991, "The Lissajous transformation. IV. Delaunay and Lissajous variables," *Celes. Mech.* **51**, 271–280.
Dieudonné, J.: 1955, *La géométrie des groupes classiques*. Springer-Verlag, Berlin - Göttingen - Heidelberg.
Dickson, L. E.: 1918, "On quaternions and their generalization and the history of the eight square problem," *Ann. Math.* **20**, 155–171.
Ferrándiz, J. M.: 1986, "Una transformación canónica general que aumenta el número de variables. Aplicación a problemas de dos cuerpos," *Homenaje al profesor D. Rafael Cid*, Ed. Vicente Camarena Badia. Universidad de Zaragoza, Secretariado de Publicaciones.
Ferrándiz, J. M.: 1986, "A new set of canonical variables for orbit calculations," Second International Symposium on Spacecraft Flight Dynamics, European Space Agency **SP-255**, 361–364.
Ferrándiz, J. M.: 1987, "Extended Canonical Transformations increasing the Number of Variables," in *Long-Term Dynamical Behaviour of Natural and Artificial N-Body Systems*, ed. A. E. Roy, Kluwer Academic Publishers, Dordrecht, 103–110.
Ferrándiz, J. M.: 1988, "A general canonical transformation increasing the number of variables with application to the two-body problem," *Celes. Mech.* **41**, 343–357.
Fock, V.: 1935, "Wasserstoffatom und nicht-Euklidische Geometrie," *Izvest. Akad. Nauk SSSR, Otdel. Techn. Nauk* **2**, 169–188.
Fock, V.: 1936, "Zur Theorie des Wasserstoffatoms," *Zeits. Phys.* **98**, 145–154.
Giacaglia, G. E. O,: 1970, "The Levi-Civita problem and the Kustaanheimo-Stiefel transformation," Applied Mechanics Research Laboratory, The University of Texas at Austin, TR **1018**, 23 pp..
Hurwitz, A.: 1923, "Über die Komposition der quadratischen Formen," *Math. Ann.* **88**, 1–25.
Iwai, T.: 1985, 'On reduction of two degrees of freedom Hamiltonian systems by an S^1 action, and '$SO_0(1,2)$ as a dynamical group," *J. Math. Phys.* **26**, 885–892.
Jezewski, D. J.: 1983, "A noncanonical analytic solution to the J_2 perturbed two-body problem," *Celes. Mech.* **30**, 343–362.
Kantor, I. L. and Solodovnikov, A. S.: 1973, *Giperkompleksnye Čisla*. Moscow, Nauka. Translation under the title *Hypercomplex numbers. An elementary introduction to algebras*. Berlin, Springer-Verlag, 1989.
Kurcheeva, I. V.: 1977, "Kustaanheimo-Stiefel regularization and nonclassical canonical transformations," *Celes. Mech.* **15**, 353–365.
Kustaanheimo, P.: 1964, "Spinor regularization of the Kepler motion," *Annales Universitatis Turkuensis* Ser. A **73**, 7 p..
Kustaanheimo, P. and Stiefel, E.: 1965, "Perturbation theory of Kepler motion based on spinor regularization," *J. Reine Angew. Math.* **218**, 204–219.
Lambert, D. and Kibler, M.: 1988, "An algebraic and geometric approach to non-bijective quadratic transformations," *J. Phys. A: Math. Gen.* **21**, 307–343.
Laplace, P. S.: 1799, *Traité de mécanique céleste*. Paris, Duprat et Bachelier. Reprinted in *Œuvres*. Paris, Imprimerie royale, 1843.
Lidov, M. L.: 1982, "Increasing the dimensions of Hamiltonian systems, the KS-transformation, and application of particular integrals," *Kosmicheskie Issledovaniya* **20**, 163–176.
Meyer, K. R.: 1984, "Scaling Hamiltonian systems," *Siam J. Math. Anal.* **15**, 877-889.
Milankovich, M.: 1939, "On the applications of vectorial elements in the computation of a planetary perturbation," *Bull. Serbian Acad. Math. Nat.* **A**, n^{o} 6.

Moser, J.: 1970, "Regularization of Kepler's problem and the averaging method on a manifold," *Commun. Pure Appl. Math.* **23**, 609–636.

Musen, P.: 1954, "Special perturbations of the vectorial elements," *A. J.* **59**, 262–267.

Musen, P.: 1961, "On Strömgren's Method of Special Perturbations," *Journal of the Astronautical Sciences* **8**, 49–51.

Palacios M., Abad, A., and Elipe, A.: 1991, "An efficient numerical formulation for orbit computation," in *AAS/AIAA Astrodynamics Specialist Conference*, Durango, CO, U.S.A., Paper AAS 91-414, 10 p..

Palacios, M. and Calvo, C.: 1992, "Quaternions and numerical orbit computation," Rev. Academia de Ciencias, Zaragoza **47**, 169–181.

Scheifele, G.: 1968, "Verallgemeinerte kanonische Transformationen zur Abbildung von Keplerbewegungen auf harmonische Oszillatoren und zugehörige Störungstheorie," *Dissertation* n° **4212**, Eidgenössische Technische Hochschule, Zurich.

Scheifele, G.: 1970, "On nonclassical canonical systems," *Celes. Mech.* **2**, 296–310.

Stiefel, E.: 1970, "Remarks on numerical integration of Keplerian orbits," *Celes. Mech.* **2**, 274–281.

Stiefel, E. and Scheifele, G.: 1968, "Régularisation du mouvement képlérien perturbé dans l'espace à trois dimensions par une transformation canonique généralisée," *C. R. Acad. Sc. Paris* **Série A 267**, 950–953.

Stiefel, E. and Scheifele, G.: 1971, *Linear and Regular Celestial Mechanics.* New York Heidelberg Berlin, Springer-Verlag.

Velte, W.: 1978, "Concerning the regularizing KS-transformation," *Celes. Mech.* **17**, 395–403.

Vivarelli, M. D.: 1983, "The KS-transformation in hypercomplex form," *Celes. Mech.* **29**, 45–50.

Vivarelli, M. D.: 1988, "Geometrical and physical outlook on the cross product of two quaternions," *Celes. Mech.* **41**, 359–370.

Vivarelli, M. D.: 1991, "On the connection among three classical mechanical problems via the hypercomplex KS-transformation," *Celes. Mech.& Dyn. Astron.* **50**, 109–124.

Wene, G. P.: 1984, "Algebras with anticommuting basal elements, space-time symmetries and quantum theory," *J. Math. Phys.* **25**, 414–416.

Wintner, A.: 1947, *The Analytical Foundations of Celestial Mechanics*, Princeton University Press.

Wolfram, S.: 1991, *Mathematica. A System for Doing Mathematics by Computer*, Second Edition. Addison-Wesley Publishing Company, Inc.