SIAM J. SCI. COMPUT.
Vol. 30, No. 4, pp. 2084–2112

THE EXACT COMPUTATION OF THE FREE RIGID BODY MOTION AND ITS USE IN SPLITTING METHODS*

E. CELLEDONI†, F. FASSÒ‡, N. SÄFSTRÖM†, AND A. ZANNA§

Abstract. This article investigates the use of the computation of the exact free rigid body motion as a component of splitting methods for rigid bodies subject to external forces. We review various matrix and quaternion representations of the solution of the free rigid body equation which involve Jacobi ellipic functions and elliptic integrals and are amenable to numerical computations. We consider implementations which are exact (i.e., computed to machine precision) and semiexact (i.e., approximated via quadrature formulas). We perform a set of extensive numerical comparisons with state-of-the-art geometrical integrators for rigid bodies, such as the preprocessed discrete Moser–Veselov method. Our numerical simulations indicate that these techniques, combined with splitting methods, *can* be favorably applied to the numerical integration of torqued rigid bodies.

Key words. rigid body, Jacobi elliptic integrals, splitting methods, attitude rotation, quaternions

AMS subject classifications. Primary, 70-08, 37M15; Secondary, 70E40, 65P10

DOI. 10.1137/070704393

1. Introduction.

The accurate and efficient integration of the equations of motion of a rigid body under the influence of conservative forces is of great interest in various fields, particularly mechanics and molecular dynamics (see, e.g., [17]). Splitting algorithms are frequently used: the Hamiltonian $H = T + V$, where T is the kinetic energy and V is the potential energy, is written as the sum of integrable terms, whose individual flows can be computed accurately and efficiently (see [11, 21] for a background on splitting methods).

If the body has two equal moments of inertia, then the flow of T, namely, the flow of the *free rigid body*, involves only trigonometric functions. Therefore splittings based on the computations of the flows of T and of V are widely used; see, e.g., [28, 7, 3]. If the body has three different moments of inertia, instead, it is common practice to further split the flow of T into a number of simpler flows, each of which is computable in terms of trigonometric functions; see [28, 20, 26, 7, 9]. However, it is a classical result which dates back to Legendre and Jacobi [13] that, even in the case of three distinct moments of inertia, the flow of the free rigid body can be explicitly integrated in terms of special functions—Jacobi elliptic functions for the angular momentum equation and elliptic integrals or theta functions for the attitude equation; see, e.g., [2, 31, 16, 14]. Hence, the flow of T is numerically computable and can be used as a component of splitting algorithms. Because of this, there has recently been a renewal of interest in the exact integration of the free rigid body and in its use in splitting methods; see, for example, [6, 29, 30].

The aim of this article is to investigate the potentialities of this approach through extended comparisons with other existing methods, particularly with those which

*Received by the editors October 2, 2007; accepted for publication (in revised form) December 20, 2007; published electronically May 14, 2008.
http://www.siam.org/journals/sisc/30-4/70439.html
†Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway (Elena.Celledoni@math.ntnu.no, Niklas.Savstrom@math.ntnu.no).
‡Dipartimento di Matematica Pura ed Applicata, Università di Padova, Via G. Belzoni 7, 35131 Padova, Italy (fasso@math.unipd.it).
§Department of Mathematics, University of Bergen, Johannes Brunsgt. 12, 5008 Bergen, Norway (Antonella.Zanna@math.uib.no).

appear to be the state of the art for the integration of the free rigid body with distinct moments of inertia, that is, a number of splitting algorithms [28, 20, 26, 7] and the so-called preprocessed discrete Moser–Veselov method of [12]. The latter method consists of applying the classical discrete Moser–Veselov algorithm [24] to a free rigid body whose moments of inertia have been suitably modified. This produces high order approximations of the solution of the original free rigid body problem. The modified moments of inertia depend on the initial conditions through the integrals of motion and are given by series expansions in powers of the time step. Truncations of these series produce integrators of arbitrarily high orders at a very moderate increase in computational cost. See also [22] for an earlier version of this approach.

The rigid body motion can be described in a variety of ways, for example, by using Euler angles, rotation matrices, and quaternions. A variety of expressions of the solution of the equations of motion has been given in each case. In section 2 we derive expressions of the solution which are amenable for numerical computations by using both rotation matrices and quaternions (as they are generally preferred to Euler angles in numerical algorithms), and we discuss the link between them. Even though this is of course nothing else than a revisitation of classical material, we add a unified and mathematically precise treatment. We also discuss the relationship to other approaches known in the literature [15, 29].

We consider the implementation of two of these algorithms, one with rotation matrices and one with quaternions. They both use the elliptic integral of the third kind, and to compute this function we consider two strategies. One is *exact*, that is, computes the required functions to machine precision by means of the well-known Carlson method [25]. The other, that we call *semiexact*, uses Gaussian quadrature of arbitrarily high order and produces high order approximations of the solution of the free rigid body. At the price of making the error in the evaluation of the integral depending on the step size of integration, this allows a substantial reduction of the computational cost.

In section 3 we perform numerical experiments. In particular, in section 3.2 we consider the free rigid body and compare the methods with those of [12]. We investigate how the different methods perform for different choices of the moments of inertia. It should be noted that, as far as the free rigid body is concerned, an obvious yet important feature of the exact methods is that they can be applied with any value of the time step, while approximate and semiexact methods must be applied with small enough time steps in order to achieve a desired accuracy. Furthermore, implicit methods, such as those of [12], use fixed-point iteration, which might require small step sizes to converge. These numerical comparisons give some indication on the potentiality of exact and semiexact algorithms as components of splitting methods for forced rigid bodies. In fact, compared to approximate methods, exact and semiexact methods are more robust in their dependence on the size of the time step. In particular, these methods perform better, compared to others, when using large step sizes.

Next, in sections 3.3 and 3.4 we investigate numerically the use of exact and semiexact methods as components of splitting methods for the integration of some problems involving rigid bodies subject to external forces. Specifically, we consider some sample cases with and without a fixed point and a case from molecular dynamics. In molecular dynamics situations the large number of particles implies that most of the computation time is spent to evaluate the interacting forces, so that an increase in the time spent to update the individual rigid molecules' state can be compensated by the advantage given by the use of larger step sizes.

Altogether, our conclusion is that the implementation of the exact solution of the free rigid body is in general a competitive approach compared to other numerical methods, which is worthy of consideration.

2. The exact solution for the free rigid body.

2.1. The equations of motion. The configuration of a rigid body with a fixed point is determined by the rotation which transforms a chosen orthonormal frame $\{\boldsymbol{E}_1^s, \boldsymbol{E}_2^s, \boldsymbol{E}_3^s\}$ fixed in space into a chosen orthonormal frame $\{\boldsymbol{E}_1^b, \boldsymbol{E}_2^b, \boldsymbol{E}_3^b\}$ attached to the body, both having the origin in the body's fixed point. We assume that $\boldsymbol{E}_1^b, \boldsymbol{E}_2^b, \boldsymbol{E}_3^b$ are principal axes of inertia of the body. As is customary, we identify all vectors with their representatives in the body base that we denote with lowercase fonts (that is, $\boldsymbol{v} = (v_1, v_2, v_3)^T$ is the body representative of $\boldsymbol{V} = \sum_i v_i \boldsymbol{E}_i^b$) and denote by $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ the vectors of the canonical basis of $\mathbb{R}^3$. The configuration of the body is thus determined by the attitude matrix $Q \in \mathrm{SO}(3)$ which transforms body representives into spatial representatives of vectors; in particular, $Q\boldsymbol{e}_i^s = \boldsymbol{e}_i$ for $i = 1, 2, 3$.

If $\boldsymbol{m} = (m_1, m_2, m_3)^T$ is the body representative of the angular momentum vector and $I = \mathrm{diag}(I_1, I_2, I_3)$ is the inertia tensor, then the equations of motion can be written as

$$\dot{\boldsymbol{m}} = \boldsymbol{m} \times I^{-1}\boldsymbol{m}, \tag{2.1}$$

$$\dot{Q} = Q\,\widehat{I^{-1}\boldsymbol{m}}. \tag{2.2}$$

Here $\times$ denotes the vector product in $\mathbb{R}^3$ and the hat-map $\widehat{\ }\colon \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined as

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad \mapsto \quad \widehat{\boldsymbol{v}} = \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix}$$

and satisfies $\widehat{\boldsymbol{v}}\boldsymbol{u} = \boldsymbol{v} \times \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$.

Equation (2.1) is the Euler equation (written for the angular momentum rather than for the angular velocity $\boldsymbol{\omega} = I^{-1}\boldsymbol{m}$), while (2.2) is sometimes called the Arnold equation. These are the left-trivialized Hamilton equations on $T^*\mathrm{SO}(3) \approx \mathrm{SO}(3) \times \mathbb{R}^3 \ni (Q, \boldsymbol{m})$, with the kinetic energy

$$T = \frac{m_1^2}{2I_1} + \frac{m_2^2}{2I_2} + \frac{m_3^2}{2I_3}$$

as the Hamiltonian. These equations form a completely integrable Hamiltonian system—in fact, a superintegrable or noncommutatively integrable system since, besides the kinetic energy, also the three components of the spatial angular momentum vector $Q\boldsymbol{m}$ are constants of motion (see, for instance, [10] and references therein). In particular, the norm of the body angular momentum $G = \|\boldsymbol{m}\|$ is a constant of motion.

As we review in this section, (2.1) and (2.2) can be explicitly integrated in terms of elliptic functions. The integration is done in two steps. First, the Euler equation (2.1) is integrated to give $\boldsymbol{m}(t)$. Then the Arnold equation (2.2) becomes a time-dependent linear equation for $Q(t)$, whose integration exploits in an essential manner the constancy of the spatial angular momentum vector. We shall review different representations of the solution, including some that use quaternions instead of rotation matrices.

Note that, due to the obvious SO(3)-symmetry and scaling invariance of (2.1) and (2.2), we may restrict ourselves to describe their solutions with initial conditions $(Q_0, \boldsymbol{m}_0)$ at $t = t_0$ such that

$$Q_0 = 1\,, \qquad \|\boldsymbol{m}_0\| = 1\,.$$

We shall indeed do so in order to keep the notational complexity to a minimum, but we shall indicate the changes which give the general solutions. Depending on notational convenience, we shall indifferently write $\boldsymbol{m}(t)$ or $\boldsymbol{m}_t$ for the value at time t of the solution of the Euler equation, etc.

From now on, we tacitly assume that the three moments of inertia I_1, I_2, and I_3 are pairwise distinct, and we order them in ascending order: $I_1 < I_2 < I_3$.

2.2. Solution of the Euler equation. The integration of the Euler equation (2.1) is a standard matter, and we restrict ourselves to provide the result. As is well known, the Euler equation can be viewed as a Hamiltonian system with respect to the Lie–Poisson structure on $\mathbb{R}^3 \approx \mathfrak{so}(3)^*$ and has the energy T and the norm of the angular momentum $G := \|\boldsymbol{m}\|$ as constants of motion. For given $G > 0$, the phase portrait consists of the six equilibria $\pm G\boldsymbol{e}_j$, $j = 1, 2, 3$, of the four stable/unstable manifolds of the equilibria $\pm G\boldsymbol{e}_2$, which are given by $2TI_2 = G^2$, and of periodic orbits which fill four disconnected regions of the sphere $G = \text{const}$. The periodic orbits satisfy either $2TI_3 > G^2 > 2TI_2$ or $2TI_2 > G^2 > 2TI_1$, and, for given T and G, there are two of them.

The expression of the periodic solutions involve the three Jacobi elliptic functions sn, cn, and dn, whose definition is recalled in Appendix A. As mentioned, we consider only solutions with unit norm. Given T, define the positive constants

$$I_{jh} = |I_j - I_h|\,, \qquad \Delta_j = |1 - 2TI_j|\,, \qquad B_{jh} = \left(\frac{I_j \Delta_h}{I_{jh}}\right)^{1/2}$$

for $j, h = 1, 2, 3$, $j \neq h$, and

$$k = \left(\frac{\Delta_1 I_{32}}{\Delta_3 I_{21}}\right)^{1/2}, \quad \lambda_1 = \left(\frac{\Delta_1 I_{23}}{I_1 I_2 I_3}\right)^{1/2}, \quad \lambda_3 = \left(\frac{\Delta_3 I_{12}}{I_1 I_2 I_3}\right)^{1/2}$$

that we shall use without reference throughout this section.

PROPOSITION 2.1. *Let $\boldsymbol{m}_t$ be a solution of the Euler equation* (2.1) *with unit norm and energy T.*

(i) *If $2TI_2 > 1 > 2TI_1$, then*

$$\boldsymbol{m}_t = \Big(\sigma B_{13}\,\mathrm{dn}(\lambda t - \nu, k)\,,\ B_{21}\,\mathrm{sn}(\lambda t - \nu, k)\,,\ B_{31}\,\mathrm{cn}(\lambda t - \nu, k)\Big)^T, \tag{2.3}$$

with $\lambda = \sigma\lambda_3$, for some $\nu \in \mathbb{R}$ and $\sigma = \pm 1$.

(ii) *If $2TI_2 < 1 < 2TI_3$, then*

$$\boldsymbol{m}_t = \Big(B_{13}\,\mathrm{cn}(\lambda t - \nu, k^{-1})\,,\ B_{23}\,\mathrm{sn}(\lambda t - \nu, k^{-1})\,,\ \sigma B_{31}\,\mathrm{dn}(\lambda t - \nu, k^{-1})\Big)^T,$$

with $\lambda = \sigma\lambda_1$, for some $\nu \in \mathbb{R}$ and $\sigma = \pm 1$.

(iii) *If $2TI_2 = 1$ and $\boldsymbol{m}_t$ is not an equilibrium solution, then*

$$\boldsymbol{m}_t = \Big(\sigma' B_{13}\,\mathrm{sech}(\lambda t - \nu)\,,\ \tanh(\lambda t - \nu)\,,\ \sigma' B_{31}\,\mathrm{sech}(\lambda t - \nu)\Big)^T,$$

with $\lambda = \sigma\lambda_3$, for some $\nu \in \mathbb{R}$, $\sigma = \pm 1$, and $\sigma' = \pm 1$.

The proof of these expressions reduces to differentiation; see, e.g., [16]. Solutions on the stable/unstable manifolds have been included mostly for completeness, as their need in numerical computations is quite rare. Note that in the first two cases the phase ν can be taken modulo the period of the Jacobi elliptic functions.

Remark. Solutions with norm G are obtained from the formulas of Proposition 2.1 with the substitutions $\boldsymbol{m} \mapsto G\boldsymbol{m}$ and $T \mapsto T/G^2$.

2.3. Integration of the rotation matrix. There are various derivations of the solution $t \mapsto Q_t$ of (2.2) for the attitude matrix. They all have in common the use of the constancy of the angular momentum vector in space to reduce the determination of Q_t to the determination of a planar rotation which, thanks to the knowledge of the solution of Euler equation, reduces to the evaluation of the integral of a known function. The procedure is more easily explained in terms of space vectors rather than of their body representatives.

Let $\boldsymbol{M}$ be the angular momentum vector that, as above, we assume of unit norm, $\mathcal{B}^s = \{\boldsymbol{E}_1^s, \boldsymbol{E}_2^s, \boldsymbol{E}_3^s\}$ be the spatial frame, and $\mathcal{B}^b = \{\boldsymbol{E}_1^b, \boldsymbol{E}_2^b, \boldsymbol{E}_3^b\}$ be the body frame. $\boldsymbol{M}$ and $\mathcal{B}^s$ are fixed in space, while $\mathcal{B}^b$ changes with time. Consider any rotation $\mathcal{P}_t$ which takes $\boldsymbol{M}$ into the position of $\boldsymbol{E}_3^b$ at time t; this rotation depends on t, and its inverse transforms the body basis $\mathcal{B}^b$ into a certain orthonormal frame $\mathcal{B}_t = \{\boldsymbol{V}_t, \boldsymbol{W}_t, \boldsymbol{M}\}$. Similarly, let $\mathcal{R}$ be a (time-independent) rotation which transforms $\boldsymbol{E}_3^s$ into $\boldsymbol{M}$ and hence the spatial basis $\mathcal{B}^s$ into a certain orthonormal frame $\mathcal{B}' = \{\boldsymbol{V}', \boldsymbol{W}', \boldsymbol{M}\}$. Since the frames $\mathcal{B}'$ and $\mathcal{B}_t$ have the $\boldsymbol{M}$ axis in common, there is a (time-dependent) rotation $\mathcal{Y}_t$ of axis $\boldsymbol{M}$ which transforms the former into the latter. Therefore, the rotation $\mathcal{Q}_t = \mathcal{R} \circ \mathcal{Y}_t \circ \mathcal{P}_t$ transforms the spatial basis into the body basis.

This procedure is not unique in that it depends on the choice of $\mathcal{P}_t$ and $\mathcal{R}$ but has the advantage that, for each such choice, the determination of $\mathcal{Q}_t$ reduces to the determination of a rotation about a known axis, that is, of an angle. Note that, if $\mathcal{Q}_t$ equals the identity at a certain time t_0, as we may and do assume, then it is possible to choose $\mathcal{R} = \mathcal{P}_{t_0}^{-1}$ and, correspondingly, $\mathcal{Y}_{t_0} = 1$.

Translated into body coordinates, this procedure leads to a representation of the attitude matrix Q_t as the product $P_{t_0}^T Y_t P_t$, with $P_t, Y_t \in \mathrm{SO}(3)$ such that

$$P_t \boldsymbol{m}_t = \boldsymbol{e}_3 \quad \text{and} \quad Y_t \boldsymbol{e}_3 = \boldsymbol{e}_3 \quad \forall\, t\,, \quad Y_{t_0} = 1\,. \tag{2.4}$$

We begin by giving an expression for the angle ψ_t of the rotation Y_t as a function of P_t. For shortness, we do it only in case (i) of Proposition 2.1.

Here and in the following we denote by a dot the Euclidean scalar product in $\mathbb{R}^3$ (and later on also in $\mathbb{R}^4$). Moreover, we use the inner product

$$\langle A, B \rangle := \frac{1}{2} \mathrm{tr}\,(A^T B)$$

on the space of 3×3 skew-symmetric matrices. Note that $\langle \widehat{u}, \widehat{v} \rangle = u \cdot v$ for all $u, v \in \mathbb{R}^3$.

PROPOSITION 2.2. *Consider a solution $\boldsymbol{m}_t$ of the Euler equation with unit norm. Let $P_t, Y_t \in \mathrm{SO}(3)$ be smooth functions which satisfy* (2.4)*, and write $Y_t = \exp(\psi_t \widehat{\boldsymbol{e}}_3)$ for some real function ψ_t. Then*

$$Q_t := P_{t_0}^T Y_t P_t \tag{2.5}$$

is the solution of (2.2) *with initial datum $Q_{t_0} = 1$ if and only if*

$$\psi_t = \int_{t_0}^{t} \big(2T + \langle \widehat{\boldsymbol{e}_3}, P_s \dot{P}_s^T \rangle\big) ds \pmod{2\pi} \tag{2.6}$$

or, equivalently, if $\boldsymbol{v}_t$ and $\boldsymbol{w}_t$ are the first two columns of P_t^T:

$$\psi_t = \int_{t_0}^{t} \big(2T + \boldsymbol{w}_s \cdot \dot{\boldsymbol{v}}_s\big) ds \quad (\mathrm{mod} 2\pi)\,. \tag{2.7}$$

Proof. Let $\boldsymbol{\omega}_t = I^{-1}\boldsymbol{m}_t$ be the angular velocity. Under hypotheses (2.4), the matrix Q_t as in (2.5) satisfies $Q_{t_0} = 1$ if and only if $\psi_{t_0} = 0$. Thus, it suffices to prove that $Q_t = P_{t_0}^T Y_t P_t$ satisfies $\dot{Q}_t = Q_t \widehat{\boldsymbol{\omega}}_t$ if and only if

$$\dot{\psi}_t = 2T + \langle \widehat{\boldsymbol{e}_3}, P_t \dot{P}_t^T \rangle\,. \tag{2.8}$$

For simplicity, we omit the indication of the dependency on t. Since $\dot{Y} = \dot{\psi} Y \widehat{\boldsymbol{e}_3}$, differentiating (2.5) gives $\dot{Q} = QP^T(\dot{\psi}\widehat{\boldsymbol{e}_3} + \dot{P}P^T)P$. Hence, $\dot{Q} = Q\widehat{\boldsymbol{\omega}}$ if and only if $\widehat{\boldsymbol{\omega}} = P^T(\dot{\psi}\widehat{\boldsymbol{e}_3} + \dot{P}P^T)P$. Since $P\widehat{\boldsymbol{u}}P^T = \widehat{P\boldsymbol{u}}$ for all $P \in \mathrm{SO}(3)$ and $\boldsymbol{u} \in \mathbb{R}^3$, this condition is equivalent to $\dot{\psi}\widehat{\boldsymbol{e}_3} = \widehat{P\boldsymbol{\omega}} - \dot{P}P^T$, namely, $\dot{\psi} = \langle \widehat{\boldsymbol{e}_3}, \widehat{P\boldsymbol{\omega}} + P\dot{P}^T \rangle$ given that the matrices $\widehat{\boldsymbol{e}_1}, \widehat{\boldsymbol{e}_2}, \widehat{\boldsymbol{e}_3}$ form an orthonormal set for the inner product $\langle\ ,\ \rangle$ and $\dot{P}P^T$ is skew-symmetric. The proof of (2.8) is concluded by observing that $\langle \widehat{\boldsymbol{e}_3}, \widehat{P\boldsymbol{\omega}} \rangle = \boldsymbol{e}_3 \cdot P\boldsymbol{\omega} = P^T\boldsymbol{e}_3 \cdot \boldsymbol{\omega} = \boldsymbol{m} \cdot \boldsymbol{\omega} = 2T$.

Let us now prove (2.7). From $P^T\boldsymbol{e}_3 = \boldsymbol{m}$ it follows that $P = [\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{m}]^T$, with orthonormal vectors $\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{m}$, and one computes $P\dot{P}^T = -\boldsymbol{w}\cdot\dot{\boldsymbol{m}}\widehat{\boldsymbol{e}_1} + \boldsymbol{v}\cdot\dot{\boldsymbol{m}}\widehat{\boldsymbol{e}_2} - \boldsymbol{v}\cdot\dot{\boldsymbol{w}}\widehat{\boldsymbol{e}_3}$. Thus $\langle \widehat{\boldsymbol{e}_3}, P\dot{P}^T \rangle = -\boldsymbol{v}\cdot\dot{\boldsymbol{w}} = \dot{\boldsymbol{v}}\cdot\boldsymbol{w}$. ∎

Note that any unit vector $\boldsymbol{v}_t$ orthogonal to $\boldsymbol{m}_t$ can be used to construct the matrix $P_t = [\boldsymbol{v}_t, \boldsymbol{w}_t, \boldsymbol{m}_t]^T$, where $\boldsymbol{w}_t = \boldsymbol{m}_t \times \boldsymbol{v}_t$. Since $\|\boldsymbol{m}_t\| = 1$ implies that $\dot{\boldsymbol{m}}_t$ is orthogonal to $\boldsymbol{m}_t$, a possible choice is that of taking $\boldsymbol{v}_t$ aligned with $\dot{\boldsymbol{m}}_t$. We specialize the expression of the angle ψ_t corresponding to this choice. For another choice, see section 2.6. The expression of ψ uses the elliptic integral of the third kind Π and the amplitude function am, whose definitions are recalled in Appendix A.

Corollary 2.3. *Consider a solution $\boldsymbol{m}_t$ of the Euler equation as in* (2.3)*, with unit norm and energy T such that $2TI_2 > 1 > 2TI_1$. If, in Proposition* 2.2*, $\boldsymbol{v}_t = \|\dot{\boldsymbol{m}}_t\|^{-1}\dot{\boldsymbol{m}}_t$, then*

$$\psi_t = 2T\,(t - t_0) + \frac{\Delta_2}{\lambda I_2}\Big[\Pi\big(\mathrm{am}(\lambda t - \nu), n, k\big) - \Pi\big(\mathrm{am}(\lambda t_0 - \nu), n, k\big)\Big], \tag{2.9}$$

with k, λ, and ν as in (2.3) *and $n = B_{23}^{-1}$.*

Proof. The orthogonality of $\boldsymbol{w} = \boldsymbol{m} \times \boldsymbol{v}$ and $\dot{\boldsymbol{m}}$ implies that $\boldsymbol{w}\cdot\dot{\boldsymbol{v}} = \boldsymbol{w}\cdot\ddot{\boldsymbol{m}}/\|\dot{\boldsymbol{m}}\|$. Since $\ddot{\boldsymbol{m}} = \frac{d}{dt}(\boldsymbol{m} \times \boldsymbol{\omega}) = \dot{\boldsymbol{m}} \times \boldsymbol{\omega} + \boldsymbol{m} \times \dot{\boldsymbol{\omega}}$ and $\dot{\boldsymbol{\omega}} = \|\dot{\boldsymbol{m}}\| I^{-1}\boldsymbol{v}$, this gives $\boldsymbol{w}\cdot\dot{\boldsymbol{v}} = \boldsymbol{v}\cdot I^{-1}\boldsymbol{v} - \boldsymbol{\omega}\cdot\boldsymbol{m} = \boldsymbol{v}\cdot I^{-1}\boldsymbol{v} - 2T$. But from Proposition 2.2 we know that $\dot{\psi} = 2T + \boldsymbol{w}\cdot\dot{\boldsymbol{v}}$. Hence $\dot{\psi} = \boldsymbol{v}\cdot I^{-1}\boldsymbol{v}$. By inserting $\dot{\boldsymbol{m}} = \boldsymbol{m} \times I^{-1}\boldsymbol{m}$ into $\boldsymbol{v}$, this becomes

$$\dot{\psi} = \frac{I_1(I_{23}m_2m_3)^2 + I_2(I_{13}m_1m_3)^2 + I_3(I_{12}m_1m_2)^2}{(I_1I_{23}m_2m_3)^2 + (I_2I_{13}m_1m_3)^2 + (I_3I_{12}m_1m_2)^2}\,.$$

By using the constancy of T and G^2 $(= 1)$ to express m_1^2 and m_3^2 in terms of T and m_2^2, and then using the expression of m_2 from (2.3), this gives

$$\dot{\psi} = 2T - \frac{I_2\Delta_1\Delta_2\Delta_3}{I_2^2\Delta_1\Delta_3 - I_{12}I_{23}m_2^2} = 2T + \frac{\Delta_2/I_2}{1 - B_{23}^{-2}\,\mathrm{sn}^2(\lambda t - \nu)}\,.$$

The proof is concluded by integrating between t_0 and t, taking into account (4.2) of Appendix A. ∎

This algorithm equals that of [16], except for the sign of ψ. A similar algorithm is given in [8].

Remark. If $2TI_3 > 1 > 2TI_2$, then ψ_t is as in (2.9) with k replaced by k^{-1}, with λ and ν as in point (ii) of Proposition 2.2, and with $n = B_{21}^{-1}$.

2.4. The equations of motion in quaternionic form. We consider now the quaternionic formulation of the free rigid body. For general references on quaternions, see, e.g., [19]. Quaternions (of unit norm) are the points of the three sphere $\mathrm{S}^3 = \{q \in \mathbb{R}^4 \,:\, \|q\| = 1\}$ equipped with a certain Lie group structure. As is customary, we write $q = (q_0, \boldsymbol{q}) \in \mathbb{R} \times \mathbb{R}^3$ and refer to q_0 and $\boldsymbol{q} = (q_1, q_2, q_3)$ as the scalar and vector parts of q, respectively. Then

$$\mathrm{S}^3 = \{q = (q_0, \boldsymbol{q}) \in \mathbb{R} \times \mathbb{R}^3 \,:\, \; q_0^2 + \|\boldsymbol{q}\|^2 = 1\}$$

is a Lie group with product

$$(p_0, \boldsymbol{p})(q_0, \boldsymbol{q}) := (p_0 q_0 - \boldsymbol{p} \cdot \boldsymbol{q}, p_0 \boldsymbol{q} + q_0 \boldsymbol{p} + \boldsymbol{p} \times \boldsymbol{q}). \tag{2.10}$$

The identity element of S^3 is $e = (1, \mathbf{0})$, and the inverse of $q = (q_0, \boldsymbol{q}) \in \mathrm{S}^3$ is $q^{-1} = (q_0, -\boldsymbol{q})$.

The "Euler–Rodriguez" map $\mathcal{E} : \mathrm{S}^3 \to \mathrm{SO}(3)$ defined by

$$\mathcal{E}(q) = 1 + 2q_0 \widehat{\boldsymbol{q}} + 2\widehat{\boldsymbol{q}}^2 \tag{2.11}$$

is a 2 : 1 surjective submersion. It is not injective since $\mathcal{E}(q) = \mathcal{E}(-q)$ and each rotation matrix has two preimages. Hence, S^3 is a double covering of $\mathrm{SO}(3)$. If $\mathcal{E}(q)$ is a rotation of angle ψ and axis $\boldsymbol{e} \in \mathbb{R}^3$, $\|\boldsymbol{e}\| = 1$, then $q = (\cos \frac{\psi}{2}, \pm \boldsymbol{e} \sin \frac{\psi}{2})$. Moreover, the map $\mathcal{E}$ is a group homomorphism since

$$\mathcal{E}(qp) = \mathcal{E}(q)\mathcal{E}(p) \quad \forall q, p \in \mathrm{S}^3.$$

Thus, the quaternionic formulation of the equations of motion of the rigid body is a formulation on a covering space. Each motion of the rigid body in $\mathrm{SO}(3)$ corresponds to two (nonintersecting) motions in S^3, and it is immaterial which one is considered. The "equation of motion of the rigid body in quaternion form" is the differential equation on $T^*\mathrm{S}^3$ which describes these motions. Analogously to what is done in the case of $\mathrm{SO}(3)$, we give this equation in left-trivialized form.

The Lie algebra $\mathfrak{s}^3 = T_e\mathrm{S}^3$ of S^3 can be identified with $\mathbb{R}^3$ equipped with the cross product as a commutator. It is convenient, however, to identify $\mathfrak{s}^3$ with the subspace $\{0\} \times \mathbb{R}^3$ of $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$:

$$\mathfrak{s}^3 = \left\{u = (0, \boldsymbol{u}) \,:\, \boldsymbol{u} \in \mathbb{R}^3\right\}$$

so as to be able to exploit the fact that the quaternion product (2.10) extends to $\mathbb{R}^4$. Note that, if $u = (0, \boldsymbol{u})$ and $v = (0, \boldsymbol{v})$ are in $\mathfrak{s}^3$, then $uv = (-\boldsymbol{u} \cdot \boldsymbol{v}, \boldsymbol{u} \times \boldsymbol{v}) \in \mathbb{R} \times \mathbb{R}^3$ need not be in $\mathfrak{s}^3$. Instead, if $u = (0, \boldsymbol{u}) \in \mathfrak{s}^3$ and $q \in \mathrm{S}^3$, then $quq^{-1} \in \mathfrak{s}^3$; see also (2.14) below. We shall also use the Euclidean product of $\mathbb{R}^4$ that we denote by a dot.

A simple calculation shows that the derivative at the identity of the covering map $\mathcal{E} : \mathrm{S}^3 \to \mathrm{SO}(3)$ is the map $\mathcal{E}_* := \mathrm{T}_e\mathcal{E} : \mathfrak{s}^3 \to \mathfrak{so}(3)$ given by

$$\mathcal{E}_*(u) = 2\widehat{\boldsymbol{u}}\,, \qquad u = (0, \boldsymbol{u}) \in \mathfrak{s}^3\,. \tag{2.12}$$

If $q_t \in \mathrm{S}^3$ and $Q_t = \mathcal{E}(q_t)$, then $q_t^{-1}\dot{q}_t \in \mathfrak{s}^3$, $Q_t^T\dot{Q}_t \in \mathfrak{so}(3)$, and

$$\mathcal{E}_*(q_t^{-1}\dot{q}_t) = Q_t^T\dot{Q}_t\,. \tag{2.13}$$

By general facts about Lie groups and covering maps, the map $\mathcal{E}_*$ is a Lie algebra isomorphism and hence intertwines the two adjoint representations, that is,

$$\mathcal{E}_*(quq^{-1}) = \mathcal{E}(q)\mathcal{E}(u)\mathcal{E}(q)^{-1} \qquad \forall\, q \in \mathrm{S}^3\,,\ u \in \mathfrak{s}^3.$$

Note that this identity (which, incidentally, can be easily verified by a direct computation) can also be written as

$$\mathcal{E}_*(quq^{-1}) = 2\widehat{\mathcal{E}(q)\boldsymbol{u}} \qquad \forall\, q \in \mathrm{S}^3\,,\ u = (0,\boldsymbol{u}) \in \mathfrak{s}^3\,. \tag{2.14}$$

As a direct consequence of (2.13) and (2.14) we can now state the rigid body equations of motion on S^3.

PROPOSITION 2.4. *Assume that $\boldsymbol{m}_t$ is a solution of the Euler equation* (2.1) *and that $q_t \in \mathrm{S}^3$ is a smooth function. Then $Q_t := \mathcal{E}(q_t)$ is a solution of the Arnold equation* (2.2) *if and only if*

$$\dot{q}_t = \frac{1}{2}q_t\omega_t, \tag{2.15}$$

with $\omega_t = (0, I^{-1}\boldsymbol{m}_t)$.

Clearly, if q_t is a solution of (2.15) for a certain $\boldsymbol{m}_t$, then so is $-q_t$, and they project onto the same rigid body motion $\mathcal{E}(q_t)$ on SO(3). The choice of the initial condition q_{t_0} unambiguously selects one of the two. Even though we need not use this fact, we note for completeness that, written on $\mathfrak{s}^3$, that is, for $m_t = (0, \boldsymbol{m}_t)$, the Euler equation becomes $\dot{m}_t = \frac{1}{2}(m_t\omega_t - \omega_t m_t)$.

2.5. Integration of the quaternion. Solutions of (2.15) can be searched in a factorized form $q_t = p_{t_0}^{-1}y_t p_t$ analogous to that of section 2.3. To this end, it is sufficient to determine p_t and y_t so that $P_t := \mathcal{E}(p_t)$ and $Y_t := \mathcal{E}(y_t)$ satisfy properties (2.4).

Since $\mathcal{E}_*$ is an isomorphism, (2.14) shows that if $p \in \mathrm{S}^3$, $u = (0,\boldsymbol{u}) \in \mathfrak{s}^3$ and $v = (0,\boldsymbol{v}) \in \mathfrak{s}^3$, then $\mathcal{E}(p)\boldsymbol{u} = \boldsymbol{v}$ if and only if $pup^{-1} = v$. Thus, if we write

$$m_t = (0,\boldsymbol{m}_t) \in \mathfrak{s}^3\,, \qquad e_j = (0,\boldsymbol{e}_j) \in \mathfrak{s}^3 \qquad (j = 1,2,3)\,,$$

we see that the analogues of conditions (2.4) are

$$p_t m_t p_t^{-1} = e_3 \quad \text{and} \quad y_t e_3 y_t^{-1} = e_3 \quad \forall\, t\,, \quad y_{t_0} = e\,. \tag{2.16}$$

(The choice $y_{t_0} = -e$ would be acceptable as well.) We can now state the analogue of the first part of Proposition 2.2.

PROPOSITION 2.5. *Consider a solution $\boldsymbol{m}_t$ of the Euler equation with unit norm. Let $p_t, y_t \in \mathrm{S}^3$ be smooth functions which satisfy* (2.16). *Then*

$$q_t := p_{t_0}^{-1}y_t p_t$$

satisfies (2.15) *and $q_{t_0} = e$ if and only if $y_t = (\cos\frac{\psi_t}{2}, \boldsymbol{e}_3 \sin\frac{\psi_t}{2})$, with*

$$\psi_t = \int_{t_0}^{t} \left(2T + 2e_3 \cdot p_s\dot{p}_s^{-1}\right) ds \qquad (\mathrm{mod}\, 2\pi)\,. \tag{2.17}$$

Proof. Define $P_t := \mathcal{E}(p_t)$ and $Y_t := \mathcal{E}(y_t)$. The latter is a rotation with axis $\boldsymbol{e}_3$ if and only if $y_t = \pm(\cos\frac{\psi_t}{2}, e_3 \sin\frac{\psi_t}{2})$ for some ψ_t, but the plus sign has to be selected in order to have $y_{t_0} = e$. Since $Y_t = \exp(\psi_t \widehat{\boldsymbol{e}}_3)$, by recalling Proposition 2.2 and observing that $q_t = p_{t_0}^{-1} y_t p_t$ is a solution of (2.15) if and only if $\mathcal{E}(q_t) = P_{t_0}^T Y_t P_t$ is a solution of (2.2), we see that all we have to prove is that the expressions (2.17) and (2.6) of the angle ψ coincide, namely, that

$$2e_3 \cdot p\dot{p}^{-1} = \langle \widehat{\boldsymbol{e}_3}, P\dot{P}^T \rangle .$$

Let $P\dot{P}^T = \widehat{\boldsymbol{a}}$, with $\boldsymbol{a} \in \mathbb{R}^3$. Then (2.12) and (2.13) together show that $p\dot{p}^{-1} = (0, \frac{1}{2}\boldsymbol{a})$. Hence $2e_3 \cdot p\dot{p}^{-1} = \boldsymbol{e}_3 \cdot \boldsymbol{a} = \langle \widehat{\boldsymbol{e}}_3, \widehat{\boldsymbol{a}} \rangle$. ∎

In order to make the previous result applicable, we need to give conditions on the quaternion p_t which ensure that it satisfies $p_t m_t p_t^{-1} = e_3$ and then to express the angle ψ_t in terms of the components of p_t. This is the content of the following lemma.

LEMMA 2.6. *Consider a solution $\boldsymbol{m} = (m_1, m_2, m_3)^T : \mathbb{R} \to \mathbb{R}^3$ of the Euler equation with unit norm and $m_3(t) \neq -1$ for all t. Then four smooth functions $p_0, p_1, p_2, p_3 : \mathbb{R} \to \mathbb{R}$ are the components of a function $p : \mathbb{R} \to \mathrm{S}^3$ which satisfies* (2.16) *if and only if*

$$p_1 = \frac{p_3 m_1 + p_0 m_2}{1 + m_3}, \qquad p_2 = \frac{p_3 m_2 - p_0 m_1}{1 + m_3}, \tag{2.18}$$

$$p_0^2 + p_3^2 = \frac{1 + m_3}{2} . \tag{2.19}$$

In that case

$$2T + 2e_3 \cdot p\dot{p}^{-1} = \frac{2T + I_3^{-1} m_3}{1 + m_3} + 4\frac{p_3\dot{p}_0 - p_0\dot{p}_3}{1 + m_3} . \tag{2.20}$$

Proof. A computation shows that the four components of $pm = (-\boldsymbol{p} \cdot \boldsymbol{m}, p_0\boldsymbol{m} + \boldsymbol{p} \times \boldsymbol{m})$ equal those of $e_3 p = (-\boldsymbol{p} \cdot \boldsymbol{e}_3, p_0\boldsymbol{e}_3 - \boldsymbol{p} \times \boldsymbol{e}_3)$ if and only p_0, p_1, p_2, p_3 satisfy (2.18). Condition (2.19) ensures that (p_0, p_1, p_2, p_3) has norm one. Next, by using (2.18) one computes

$$e_3 \cdot p\dot{p}^{-1} = (p_3\dot{p}_0 - p_0\dot{p}_3) + (p_2\dot{p}_1 - p_1\dot{p}_2) = 2\frac{p_3\dot{p}_0 - p_0\dot{p}_3}{1 + m_3} - \frac{m_1\dot{m}_2 - m_2\dot{m}_1}{2(1 + m_3)},$$

and the conclusion follows by observing that $m_1\dot{m}_2 - m_2\dot{m}_1 = \boldsymbol{e}_3 \cdot \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{\omega}) = 2Tm_3 - \omega_3$, where as usual $\boldsymbol{\omega}$ is the angular velocity. ∎

Thus, any choice of p_0 and p_3 satisfying (2.19) leads to a quaternionic implementation of the free rigid body motion. For instance, taking $p_0 = c_0\sqrt{1 + m_3}$ and $p_3 = c_3\sqrt{1 + m_3}$ with constants c_0 and c_3 such that $c_0^2 + c_3^2 = \frac{1}{2}$ leads to a particularly simple expression for $\dot{\psi}$. Taking, for instance, $c_0 = \frac{1}{\sqrt{2}}$ and $c_3 = 0$ gives the following.

COROLLARY 2.7. *Consider a solution $\boldsymbol{m}(t)$ of the Euler equation as in* (2.3)*, with unit norm and energy T such that $2TI_2 > 1 > 2TI_1$. Then quaternions $p(t)$ and $y(t) = \big(\cos\frac{\psi(t)}{2}, \boldsymbol{e}_3 \sin\frac{\psi(t)}{2}\big)$ as in Proposition* 2.5 *are given by*

$$p(t) = \frac{1}{\sqrt{2}} \left(\sqrt{1 + m_3(t)}\,,\ \frac{m_2(t)}{\sqrt{1 + m_3(t)}}\,,\ -\frac{m_1(t)}{\sqrt{1 + m_3(t)}}\,,\ 0 \right),$$

$$\psi(t) = \frac{t - t_0}{I_3} + \frac{I_{31}}{I_1 I_3 \lambda}\big[\Pi\left(\varphi(t), n, k\right) + f(t) - \Pi\left(\varphi(t_0), n, k\right) - f(t_0)\big],$$

where $\varphi(s) = \mathrm{am}(\lambda s - \nu, k)$*, with* λ*,* k*, and* ν *as in* (2.3)*,* $n = -(B_{31}/B_{13})^2$*, and*

$$f(s) = B_{21}^{-1} B_{13} B_{31} \arctan\left(B_{13}^{-1} B_{21}\,\mathrm{sd}(\lambda s - \nu, k)\right).$$

Proof. If $2TI_2 > 1 > 2TI_1$, then $m_3 > -1$ for all times. With the given choice of p_0 and p_3 the right-hand side of (2.20) reduces to $\frac{2T+m_3/I_3}{1+m_3}$, namely, $\frac{1}{I_3} + \frac{\Delta_3/I_3}{1+m_3}$. From (2.3), $m_3 = a\,\mathrm{cn}(\lambda t - \nu, k)$, with $a = B_{31}$. Since $0 < a < 1$, $n := \frac{a^2}{a^2-1} < 0$ and thus [5, page 215]

$$\int \frac{du}{1 + a\,\mathrm{cn}(u,k)} = \frac{1}{1-a^2}\big[\Pi\big(\mathrm{am}(u,k), n, k\big) - a f_1(u)\big],$$

with $f_1(u) = C \tan^{-1}(C^{-1}\mathrm{sd}(u,k))$, $C = [(1-a^2)/(k^2 + (1-k^2)a^2)]^{1/2}$. The proof is concluded with a little bit of algebra. ∎

This is a rescaled version of the algorithm presented by Kosenko in [15]. This is the algorithm we use in the numerical work of the next section.

2.6. Relation between quaternion and matrix algorithms. We discuss now very shortly how to translate into quaternionic form $q = p_{t_0}^{-1} y_t p_t$ a given representation $Q_t = P_{t_0}^T Y_t P_t$ of the attitude matrix as in Proposition 2.2. This clearly reduces to determining a quaternion p_t such that $P_t = \mathcal{E}(p_t)$. This operation involves "inverting" a two-to-one map and can of course be done only up to the overall sign of p, but this is immaterial in the present context given that the product $p_{t_0}^{-1} y_t p_t$ is independent of the sign of p.

As usual, we assume that $\|\boldsymbol{m}\| = 1$ and $2TI_2 > 1 > 2TI_1$. Thus $m_3 \neq \pm 1$, and we can invoke Lemma 2.6, which implies that a quaternion p such that $\mathcal{E}(p) = P$ is determined, up to the sign, once p_3^2 and the relative signs of p_0 and p_3 are known. If $p = (p_0, p_1, p_2, p_3)$, then, from (2.11),

$$\mathcal{E}(p) = \begin{pmatrix} 1 - 2(p_2^2 + p_3^2) & -2p_0p_3 + 2p_1p_2 & 2p_0p_2 + 2p_1p_3 \\ 2p_0p_3 + 2p_1p_2 & 1 - 2(p_1^2 + p_3^2) & -2p_0p_1 + 2p_2p_3 \\ -2p_0p_2 + 2p_1p_3 & 2p_0p_1 + 2p_2p_3 & 1 - 2(p_1^2 + p_2^2) \end{pmatrix}.$$

Equating the three diagonals entries of this matrix to those of $P = [\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{m}]^T$ gives $4p_1^2 = 1 + v_1 - w_2 - m_3$, $4p_2^2 = 1 - v_1 + w_2 - m_3$, and

$$4p_3^2 = 1 - v_1 - w_2 + m_3. \tag{2.21}$$

If p_0 and p_3 are both nonzero, then their relative sign is determined by the equality

$$4p_0p_3 = v_2 - w_1,$$

which is obtained by equating entries $(1,2)$ and $(2,1)$ of the two matrices $\mathcal{E}(p)$ and P. As an example, the algorithm of Corollary 2.3 uses $\boldsymbol{v} = \|\dot{\boldsymbol{m}}\|^{-1}\dot{\boldsymbol{m}} = \|\dot{\boldsymbol{m}}\|^{-1}\boldsymbol{m} \times I^{-1}\boldsymbol{m}$ and hence $\boldsymbol{w} = \boldsymbol{m} \times \boldsymbol{v} = \|\dot{\boldsymbol{m}}\|^{-1}(2T\boldsymbol{m} - I^{-1}\boldsymbol{m})$. Thus, (2.21) gives

$$p_3^2 = \frac{1+m_3}{4} + \frac{I_{32}m_2}{4I_2I_3\|\dot{\boldsymbol{m}}\|}\left(m_3 - B_{32}\right).$$

The other components of the quaternion p are computed as just explained, and the angle ψ is as in Corollary 2.3.

As another example, take $\boldsymbol{v} = \frac{\boldsymbol{m}\times\boldsymbol{e}_3}{\|\boldsymbol{m}\times\boldsymbol{e}_3\|}$. Then $v_1 = (1-m_3^2)^{-1/2}m_2$, $w_2 = (1-m_3^2)^{-1/2}m_2m_3$, and

$$p_3^2 = \frac{1+m_3}{4} - \frac{m_2(1+m_3)}{4\sqrt{1-m_3^2}}\,.$$

This produces a quaternion version of the algorithm based on rotation matrices recently considered by van Zon and Schofield [29]. The rotation angle is

$$\psi = \int_{t_0}^{t} \frac{2TI_3 - m_3^2}{I_3(1-m_3^2)}ds = \frac{t-t_0}{I_3} + \frac{I_{31}}{\lambda I_3 I_1}\big(\Pi(\mathrm{am}(\lambda t-\nu),n,k) - \Pi(\mathrm{am}(\lambda t_0-\nu,n,k))\big),$$

with $n = -B_{31}^{-2}B_{13}^{-2}$.

Remark. There is another possibility for constructing a quaternion p such that $pmp^{-1} = e_3$, which is used in [15]. This is based on the fact that, given any three orthonormal vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in \mathbb{R}^3$ and a vector $\boldsymbol{m} \in \mathbb{R}^3$ with unit norm, one has

$$pmp^{-1} = v_3, \quad \text{with} \quad p = \frac{v_2 + v_1 m}{\|v_2 + v_1 m\|}, \tag{2.22}$$

where, as usual, $v_i = (0, \boldsymbol{v}_i)$ and $n = (0, \boldsymbol{m})$. Reference [15] uses $\boldsymbol{v}_3 = \boldsymbol{e}_3$, $\boldsymbol{v}_1 = \gamma_1\boldsymbol{e}_1 + \gamma_2\boldsymbol{e}_2$, and $\boldsymbol{v}_2 = \gamma_2\boldsymbol{e}_1 - \gamma_1\boldsymbol{e}_2$, with $\gamma_1, \gamma_2 \in \mathbb{R}$. It is elementary to verify (2.22) by direct computation if $\boldsymbol{v}_i = \boldsymbol{e}_i$, $i = 1,2,3$. Otherwise, there is a quaternion $s \in \mathrm{S}^3$ such that $\mathcal{E}(s)\boldsymbol{e}_i = \boldsymbol{v}_i$, $i = 1,2,3$. Then $p = k(se_2s^{-1} + se_2s^{-1}m)$, with $k = \|se_2s^{-1}+se_2s^{-1}m\|^{-1}$, and a simple computation shows that $v_3p - pm = s\big[e_3(e_2+e_1n) - (e_2+e_1n)n\big]s^{-1}$ for $n = sms^{-1}$. Here the term between square brackets vanishes by virtue of the previous observation.

3. Numerical experiments.

3.1. Numerical implementation. The exact algorithms described in this paper require the computation of elliptic integrals of the first and third kinds. Elliptic integrals of the first kind are computed very rapidly by using standard algorithms such as the arithmetic geometric mean and ascending/descending Landen transformations [1]. These can be used also for the elliptic integral of the third kind, but their performance is not so uniform, and other algorithms are preferred instead. In [29] the authors use a method based on theta functions. Our implementation makes use of Carlson's algorithms `rf`, `rj`, and `rc` that have been acclaimed to produce accurate values for large sets of parameters. These methods are described in detail in [25] and are the most common routines for elliptic integrals of the third kind in several scientific libraries.

As mentioned in the introduction, an alternative to the exact computation of the elliptic integral of the third kind is the approximation by a quadrature method. We will refer to the methods obtained in this manner as *semiexact* methods. These, by construction, integrate the angular momentum exactly. They also preserve $Q\boldsymbol{m}$ (because of the properties of the matrix P in Proposition 2.2). Moreover, they will be time-symmetric if the underlying quadrature formula is symmetric.

In [30], the integral

$$\int_{u_0}^{u} \frac{ds}{1 - n\,\mathrm{sn}^2\, s}$$

is approximated by a quadrature based on Hermite interpolation, as the function sn and its derivative can be easily computed at the end points of the interval. Alternatively, one can write the same integral in the Legendre form:

$$\int_{\mathrm{am}(u_0)}^{\mathrm{am}(u)} \frac{d\theta}{(1 - n\sin^2\theta)\sqrt{1 - k^2\sin^2\theta}}. \tag{3.1}$$

This format is convenient when using quadrature formulas because it requires tabulating the sine function in the quadrature nodes instead of $\mathrm{sn}(\lambda(t-\nu))$. Thus, (3.1) can be approximated as

$$\int_{\mathrm{am}(u_0)}^{\mathrm{am}(u)} f(\theta)d\theta \approx \sum_{i=1}^{p} b_i f(\varphi_0 + a_i\Delta\varphi),$$

where $\Delta\varphi = \mathrm{am}(u) - \mathrm{am}(u_0)$ and b_i and a_i are weights and nodes of a quadrature formula, respectively. We use Gaussian quadrature (i.e., quadrature based on orthogonal polynomials), because of its high order. In particular, Gauss–Legendre quadrature with p points attains the maximal quadrature order $2p$. The coefficients and weights for the Gaussian–Legendre quadrature method of order 10 (5 nodes) used in this paper are reported in Appendix B. Our numerical experiments indicate that this approximation is very effective. For instance, the 5-point Gauss–Legendre quadrature (order 10) gives very accurate results even for moderately large step sizes and reduces the overall cost of the methods by 2/3.

With respect to the exact methods, the semiexact methods obtained with this approach are more directly comparable to the methods of [12].

3.2. Free rigid body. In this section we compare the algorithms discussed in this paper with the preprocessed Rattle/discrete Moser–Veselov (DMV) of Hairer and Vilmart [12]. The latter are, in our opinion, the state-of-the-art approximation methods insofar as the rigid body is concerned. The comparison is done by using FORTRAN routines. The methods we compare are `dmv6`, `dmv8`, and `dmv10`, the methods based on the preprocessed Rattle algorithms of order 6, 8, and 10, respectively, the two exact methods with the rotation matrix of section 2.3, and the rotation quaternion of section 2.5 along with their semiexact variants in which the elliptic integral is approximated by Gauss–Legendre quadrature formulas of order 6, 8, and 10. As explained in the introduction, in order to do these comparisons, we integrate the flow of the free rigid body in a time interval $[0, T_{\mathrm{fin}}]$ by repeated application of the time-h algorithms.

In the first experiment (see Figure 3.1, top plot) we display

$$\mathrm{average}_{I,\boldsymbol{m}_0} \log_{10} ||Q_{\mathrm{reference}}(T_{\mathrm{fin}}; I, \boldsymbol{m}_0) - Q_{\mathrm{computed}}(T_{\mathrm{fin}}|h; I, \boldsymbol{m}_0)||_\infty \tag{3.2}$$

(or the analogous quantity of the quaternion) against the CPU-time averages of the different methods when $T_{\mathrm{fin}} = 10$, with twenty different step sizes h ranging from about 0.4 down to 0.01. The absolute value of the indicator (3.2) corresponds to the average number of significant digits of the attitude matrix with step size h at time

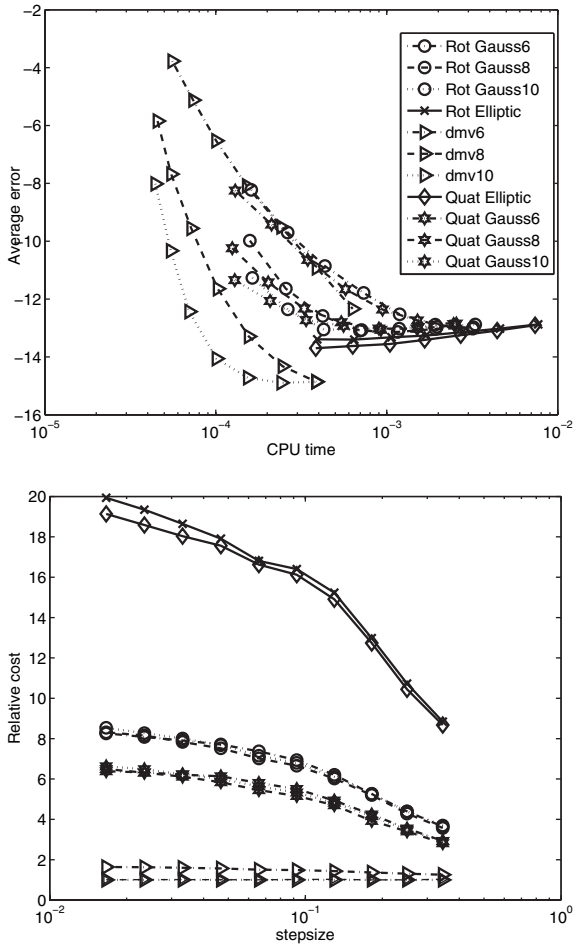

FIG. 3.1. *Left: Average log of error versus average CPU times in the attitude rotation (*100 *runs) for random initial conditions and random moments of inertia. Right: Relative cost (with respect to the cheapest method) versus step size. The methods computing the exact solutions are more expensive than the approximated ones, but their relative cost rate improves for large time steps. The DMV methods converge* 75 *out of the* 100 *runs. The failures are not taken into account when computing the averages.*

T_{fin}.[1] The set of initial parameters, shared by all of the methods, is determined as follows. We choose a random inertia tensor, normalized so that $I_1 < I_2 < I_3 = 1$, thereafter a random initial normalized angular momentum in the first quadrant. This is not a restriction, as scaling both the inertia tensor and the angular momentum is equivalent to a time reparametrization. The initial condition for the attitude matrix is the identity matrix that, for quaternions, is $(1, 0, 0, 0)$. The reference solution is computed with Matlab's `ode45` routine, setting both relative and absolute errors to machine precision. The average CPU is computed as the mean of 100 runs.

[1] Our methods compute exactly the angular momentum, while the DMV methods do not. However, both classes of methods preserve exactly the kinetic energy, the norm of the angular momentum, and the spatial angular momentum $Q\boldsymbol{m}$ and are time-reversible and Lie–Poisson integrators for the angular momentum. The DMV methods and the semiexact methods are not symplectic.

Figure 3.1 indicates that the exact methods are clearly more expensive, but they always converge (against 75 successes for the methods `dmv6`, `dmv8`, and `dmv10` that depend on a step size "small enough" for the fixed-point iterations to converge). The diverging runs of the DMV methods are not taken into account when computing averages. Good behavior is displayed also by the semiexact methods. Their cost is about 1/3 of the methods using the exact elliptic integral (and this is reasonable, because the exact routines compute 3 elliptic integrals of the third kind: the complete one between 0 and $\pi/2$ and two incomplete ones between 0 and ϕ, where $0 \le \phi \le \pi/2$). The right plot in Figure 3.1 displays the relative cost of the methods, computed as

$$\frac{\text{average cost of method X}}{\min_{\text{all methods}} \text{aver. cost of method X}},$$

so that the bottom line equals one by definition. The DMV are the cheapest methods, and their cost increases very slowly with the order of the method. We see that the relative cost of the exact and semiexact methods is higher for small step sizes and lower for large step sizes. This indicates that the exact and semiexact methods are of interest in numerical simulations that use large step sizes, for which the DMV might have problems in converging.

The exact and semiexact methods discussed in this paper reveal a worse accumulation of roundoff error for small step sizes (see Figure 3.1, left plot). This can be partly explained by the fact that the routines for the attitude rotation make repeated use of the exact solution of the angular momentum. However, with exact methods it is not necessary to perform many tiny steps for integrating to the final time: a single time stepping is enough, and this avoids the problems related to the accumulation of roundoff error. To improve on propagation of roundoff, it is also possible to perform a simple projection at the end of each time step; in the quaternion case this amounts to dividing the attitude quaternion by its norm. In general, when exact and semiexact methods are applied within a splitting method, the value of the parameters (angular momentum, attitude, energy) will change before and after one free rigid body step, and hence we do not foresee problems of roundoff accumulation.

What about the accuracy of the exact methods using matrices or quaternions? Numerical experiments reveal that the accuracy of the two exact methods is very comparable and also their cost. Methods using quaternions to represent rotations are usually faster than their matrix counterpart, but here the computational time is dominated by the evaluation of the elliptic integrals.

Our extensive numerical experiments revealed that the performance of the semiexact and the DMV methods depended heavily on the matrix of inertia I and the initial condition $\boldsymbol{m}_0$ for the angular momentum. To understand this dependence, we have followed a procedure similar to the one used in [9]. Since normalizing the matrix of inertia is equivalent to a time reparametrization, it is sufficient to consider values of the form $I_1/I_3 < I_2/I_3 < 1$. This reduces to considering two parameters, say, $x = I_1/I_3$ and $y = I_2/I_3$. As $I_i + I_j \ge I_k$, the problem is reduced to considering values of x and y in the triangle

$$\mathcal{T} = \{(x, y) \in \mathbb{R}^2 : 0 < 1 - y \le x < y < 1\}$$

(see Figure 3.2).

We construct a discretization of this triangle by superimposing a rectangular grid (100 points in the x direction and 50 in the y direction). For each point (x, y) in the interior of the triangle, we solve 20 initial value problems with initial condition $\boldsymbol{m}_0$

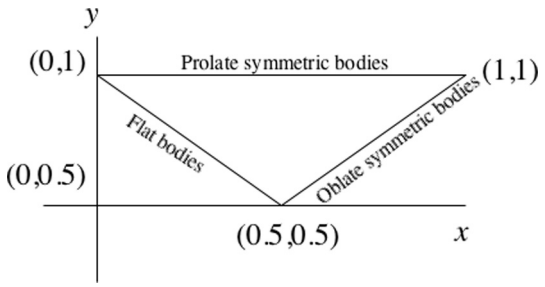

FIG. 3.2. *Parametrization domain for the matrix of inertia. x-axis:* I_1/I_3*, y-axis:* I_2/I_3*.*

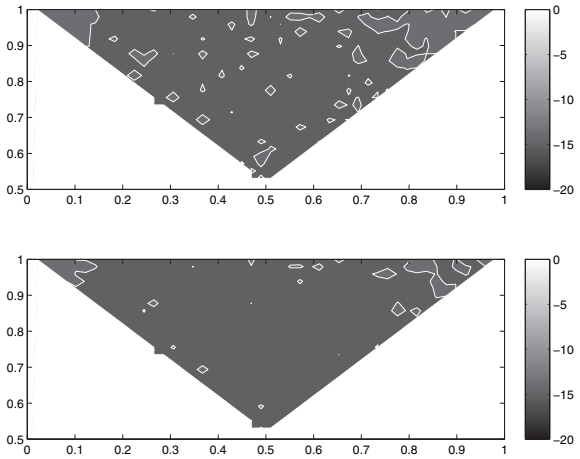

FIG. 3.3. *Average* $\log_{10}$ *error for the various values of the matrix of inertia with step size* $h = 0.4$*. Comparison of exact methods. Top: Matrix case. Bottom: Quaternion case.*

in the first octant. This set of initial parameters is identical for all of the methods. Thereafter, we compute the average (3.2) for each method (nonconverging runs for the DMV methods are discarded). The results of the experiments are shown in Figures 3.3, 3.4(a), 3.4(c), and 3.4(e), computed with integration step size $h = 0.4$, and Figures 3.4(b), 3.4(d), and 3.4(f), computed with integration step size $h = 0.04$.

For the largest step size $h = 0.4$ the exact methods described in this paper perform very similarly and show a uniform accuracy. We compare then the semiexact methods of order 6, 8, and 10 and the DMV ones of the same order. It should be mentioned that the pictures corresponding to semiexact methods using matrix rotations or quaternions are virtually indistinguishable from each order; for this reason, we show only one of the two. Both the semiexact and the DMV methods reveal a worse approximation in the proximity of the top left corner

$$0 \approx x = \frac{I_1}{I_3} \ll y = \frac{I_2}{I_3} \approx 1 = \frac{I_3}{I_3}, \tag{3.3}$$

namely, when the smallest moment of inertia is much smaller than the two others. This behavior of the numerical methods is due to the fact that, when I_1 goes to zero, one of the periods of the free rigid body motion tends to zero. To resolve these motions accurately, numerical integrators must use small step sizes. The DMV methods have on average less accuracy, and they failed to converge for several initial conditions.

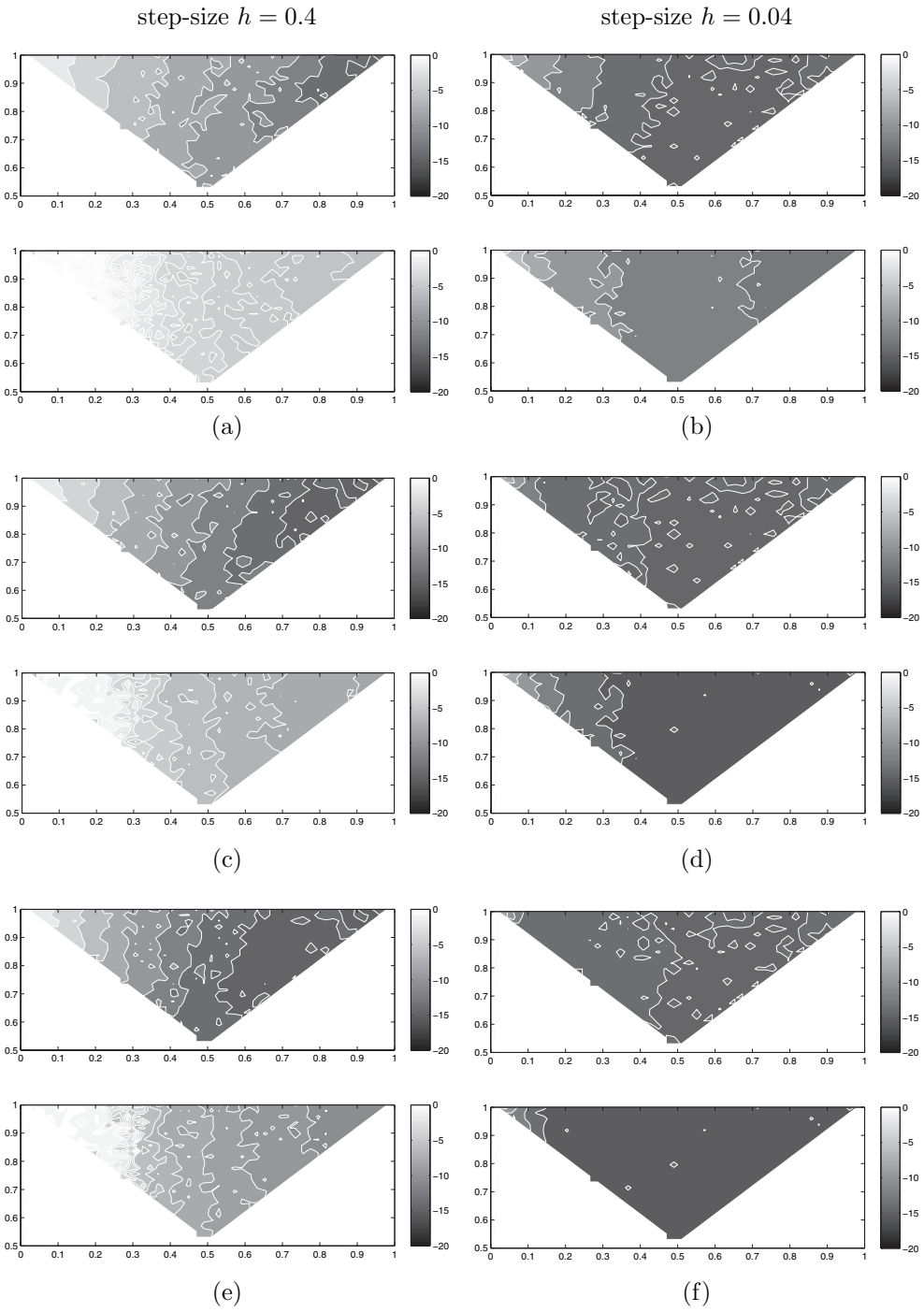

FIG. 3.4. *Average* $\log_{10}$ *error for the various values of the matrix of inertia with step size* $h = 0.4$ *(left) and* $h = 0.04$ *(right):* (a) *and* (b) *Order* 6. *Top: Semiexact with quadrature order* 6. *Bottom:* `dmv6`. (c) *and* (d) *Order* 8. *Top: Semiexact with quadrature order* 8. *Bottom:* `dmv8`. (e) *and* (f) *Order* 10. *Top: Semiexact with quadrature order* 10. *Bottom:* `dmv10`.

For the next value of the step size ($h = 0.04$) the exact methods reveal a worse accumulation of roundoff error (not shown), already observed in Figure 3.1. This accumulation disappears if the integration in $[0, T_{\text{fin}}]$ is performed with a single time step. The DMVs, in particular `dmv10`, perform very well in the whole triangle, except for the top left corner.

The conclusion is that exact and semiexact methods are of interest for large step sizes and, in particular, for values of the moments of inertia in the region (3.3).

3.3. Torqued systems and perturbations of free rigid body motions. In this section we consider systems of the form

$$H(\boldsymbol{m}, Q) = T(\boldsymbol{m}) + V(Q), \tag{3.4}$$

where T is the kinetic energy of the free rigid body and the potential energy V describes some external torque. As mentioned in the introduction, a standard approach to solve this problem is to split it into a free rigid body motion coming from the kinetic part

$$S_1 = \begin{cases} \dot{\boldsymbol{m}} = \boldsymbol{m} \times I^{-1}\boldsymbol{m}, \\ \dot{Q} = Q\,\widehat{I^{-1}\boldsymbol{m}} \end{cases} \tag{3.5}$$

plus a torqued motion, namely,

$$S_2 = \begin{cases} \dot{\boldsymbol{m}} = \mathbf{f}(Q), \\ \dot{Q} = 0, \end{cases} \tag{3.6}$$

where $\mathbf{f}(Q) = -\text{rot}(Q^T \frac{\partial V}{\partial Q})$. Here the rot function maps matrices to vectors, first by associating to a matrix a skew-symmetric one and then by identifying the latter with a vector

$$\text{rot}(A) = \text{skew}^{-1}(A - A^T),$$

where $\text{skew}(\mathbf{v}) = \hat{\mathbf{v}}$; see also [26].

Thereafter, the flows of the S_1 and S_2 systems are composed by means of a splitting method [21].

The most commonly used is the symplectic second order Störmer/Verlet scheme

$$(\boldsymbol{m}, Q)^{(j+1)} = \varphi_{h/2}^{[S_2]} \circ \varphi_h^{[S_1]} \circ \varphi_{h/2}^{[S_2]}((\boldsymbol{m}, Q)^{(j)}), \quad j = 0, 1, \ldots,$$

where $\varphi_h^{[S_1]}$ and $\varphi_h^{[S_2]}$ represent the flows of S_1 and S_2, respectively. Some higher order splitting schemes are presented in Appendix C. These are state-of-the-art optimized schemes with a very small leading error [4]. We will use these methods for the remaining experiments. All of the remaining experiments are performed in MATLAB. For the rigid body part, we use the rotation-matrix exact method of section 2.3, which we will call RB for reference.

One of the most popular methods for approximating the free rigid body system (3.5) is a second order method proposed by McLachlan and Reich (see [7]). This method, which we will call MR, is time-reversible and preserves the Poisson structure of the system. In brief, the MR method is based on a splitting of the Hamiltonian (3.4) into four parts:

$$\tilde{H}_1 = \frac{m_1^2}{2I_1}, \quad \tilde{H}_2 = \frac{m_2^2}{2I_2}, \quad \tilde{H}_3 = \frac{m_3^2}{2I_3}, \quad \tilde{H}_4 = V(Q).$$

Each of the corresponding Hamiltonian vector fields can be integrated exactly ($\tilde{H}_1$, $\tilde{H}_2$, and $\tilde{H}_3$ correspond to the vector field (3.5)), and the symmetric composition of the flows gives rise to the approximation scheme

$$(\boldsymbol{m}, Q)^{(j+1)} = \Phi_{MR}((\boldsymbol{m}, Q)^{(j)}),$$

where

$$\Phi_{MR} = \varphi_{4,h/2} \circ \Phi_{T,h} \circ \varphi_{4,h/2}.$$

Here

$$\Phi_{T,h} = \varphi_{1,h/2} \circ \varphi_{2,h/2} \circ \varphi_{3,h} \circ \varphi_{2,h/2} \circ \varphi_{1,h/2}$$

is the contribution from the kinetic parts $\tilde{H}_1$, $\tilde{H}_2$, and $\tilde{H}_3$, where the flows of the kinetic parts corresponds to elementary rotations in $\mathbb{R}^3$.

3.3.1. The heavy top. As a first study case, we consider a nearly integrable situation, the rigid body with a fixed point in a small constant-gravity field. The Hamiltonian is

$$H = T + \varepsilon V(Q), \qquad 0 < \varepsilon \ll 1, \tag{3.7}$$

with

$$V(Q) = \boldsymbol{e}_3^T Q^T \boldsymbol{u}_0$$

for a constant vector $\boldsymbol{u}_0$. The vector $\boldsymbol{u} = Q^T \boldsymbol{u}_0$ describes the position of the center of mass times the (normalized) acceleration of gravity. This potential V corresponds to $\mathbf{f}(Q) = (u_2, -u_1, 0)^T$, where u_1 and u_2 are components of $\boldsymbol{u}$.

A symplectic splitting method of order p that treats the free rigid body part exactly would typically have a nearby Hamiltonian of the form

$$\tilde{H} = H + \varepsilon V + \mathcal{O}(\varepsilon h^p),$$

and hence, if the step size of integration is small enough, the numerical error remains smaller with respect to the perturbation parameter; see, e.g., [3]. If the rigid body part is resolved by a symplectic method of order r, typically $r \geq p$, the nearby Hamiltonian has the form

$$\tilde{H} = H + \varepsilon V + \mathcal{O}(h^r) + \mathcal{O}(\varepsilon h^p),$$

and thus, in order to have an error that goes to zero as ε goes to zero, one has to take smaller step sizes h.

This behavior is displayed in Figure 3.5 for two values of ε (left plot: $\varepsilon = 10^{-3}$; right plot: $\varepsilon = 10^{-6}$). We compare different splitting schemes of various order for the system $S_1 + S_2$. Moreover, we compare the same splitting techniques using an exact method or a further MR splitting for the free rigid body motion. As the MR method has only order two, we boost its order to p (the same as the underlying splitting scheme) by using Yoshida's technique [32].

The initial conditions, identical for all of the methods, are chosen as follows. Having fixed a value of ε, we choose a random inertia tensor, normalized so that $I_1 = 1$. Having chosen the first two components of $\boldsymbol{m}_0$ randomly, the remaining one

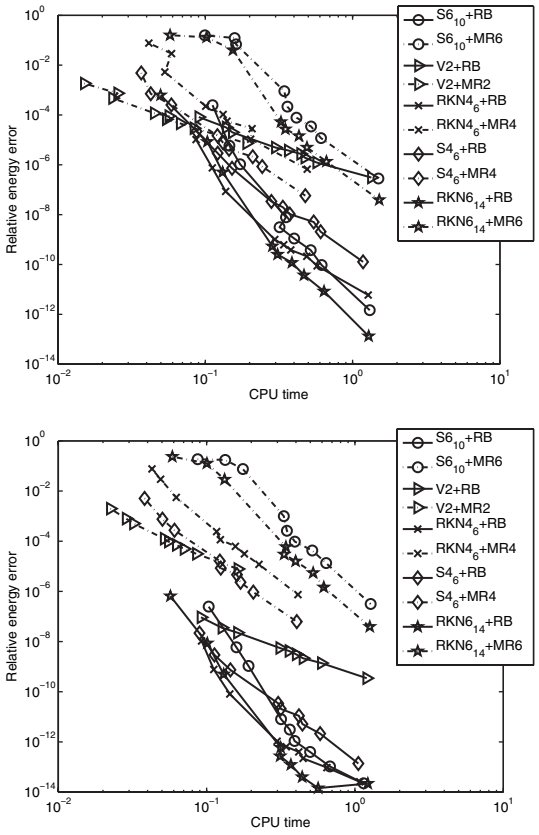

FIG. 3.5. *Average relative energy errors versus computational time, perturbed rigid body,* $\varepsilon = 10^{-3}$ *(left plot), and* $\varepsilon = 10^{-6}$ *(right plot). Initial kinetic energy* $T_0 = 1$. *Solid lines: Splitting methods using RB. Dashed-dotted lines: Splitting methods using MR approximation for the rigid body motion boosted to the same order of the splitting scheme.*

is determined to match $T_0 = 1$. The vector $\boldsymbol{u}_0$ is taken equal to $\boldsymbol{e}_3$, and Q_0 is the identity matrix.

Several splitting methods are compared, and each timing and relative Hamiltonian error is averaged (mean value) over 20 different initial conditions (each with new $I, \boldsymbol{m}_0$). The methods are implemented so that all of the splitting schemes perform the same number of force evaluations. This is done as follows: start with the following basic time steps: $h \in \{8, 5, 4, 2, 1.75, 1.5, 1.25, 1, 0.5\}$. For a splitting method with s stages (s is the number of evaluations of the force), we use $h_s = c_s h = \frac{s}{10} h$. For instance, for the sixth order 10-stages method $S6_{10}$, $c_s = 1$; for the Störmer–Verlet splitting (V2), $c_s = \frac{1}{10}$. The integration is performed in the interval $[0, 20]$.

Figure 3.5 indicates that the more we boost the order of the MR scheme, the more the cost of the splitting method becomes similar to the one using the exact solution of the rigid body. This is evident especially for schemes that have a large number of stages ($S6_{10}$, $RKN6^a_{14}$). Moreover, it is also evident that composing MR to a higher order scheme by using Yoshida's technique yields methods with a high leading error term that dominates the small error of the optimized splitting scheme. Finally, note that only the methods using the exact integrator produce an error that is smaller than ε even for very large choices of the step size. This is evident for $\varepsilon = 10^{-3}$ but,

in particular, for $\varepsilon = 10^{-6}$. The conclusion is that the use of the exact algorithm for the rigid body is definitively of interest in the integration of perturbed systems (see also [3, 6]).

3.3.2. Satellite simulation. We consider a simplified model describing the motion of a satellite in a circular orbit of radius r around the Earth [18]. Denote that $\mu = gM$, where g is the gravitational constant and M is the mass of the Earth. The potential energy of this system is given by

$$V(Q) = 3\frac{\mu}{2r^3}(Q^T\mathbf{e}_3)\cdot IQ^T\mathbf{e}_3, \tag{3.8}$$

where I is the inertia tensor and $\mathbf{e}_3$ is the canonical vector $(0,0,1)^T$ in $\mathbb{R}^3$. The torque associated to this potential becomes

$$\mathbf{f}(Q) = 3\frac{\mu}{r^3}(Q^T\mathbf{e}_3)\times I(Q^T\mathbf{e}_3). \tag{3.9}$$

We simulate the motion of the satellite by using the same parameters as in [23], namely,

$$I_1 = 1.7\times 10^4, \qquad I_2 = 3.7\times 10^4, \qquad I_3 = 5.4\times 10^4,$$

with

$$\mu = 3.986\times 10^{14}, \qquad r = 1.5\times 10^5,$$

in the interval $[0, 400]$. The initial angular momentum is $\boldsymbol{m}_0 = I\,(15,-15,15)^T$. The initial attitude Q_0 is the identity matrix. The system has an energy $H_0 = 1.21595664\times 10^7$, which is conserved in time. This experiment was also considered in [6]. The splitting method based on the exact approximation of the rigid body is very accurate. The motion of the center of mass (left column) and the relative error on the energy H_0 (right column) for the splitting method RKN6$_{14}^a$ employing our exact solution are shown in Figure 3.6. The integration is performed in the interval $[0, 400]$ with step size $h = 0.1$ (top) and $h = 0.05$ (bottom). The relative error on the energy (see Figure 3.6), which is of the order of 10^{-7} for $h = 0.1$ and 10^{-10} for $h = 0.05$, indicates that H_0 is preserved to 7 and 10 digits, respectively. The corresponding plots for the evaluation of the flow of T with the MR splitting method are shown in Figure 3.7.

3.4. Molecular dynamics simulation: Soft dipolar spheres. We consider a molecular dynamics simulation, where molecules are modeled as dipolar soft spheres. This model is of interest because it can be used to study water and aqueous solutions, as water molecules can be described by small dipoles. We consider the system described in example b in Appendix A of [7], which we recall here for completeness. Denote by m_i the total mass of the ith body, by $\mathbf{q}_i$ the position of its center of mass, by $\mathbf{p}_i$ its linear momentum, by $\mathbf{Q}_i$ its orientation, and, finally, by $\boldsymbol{m}_i$ its angular momentum in the body frame. The system has Hamiltonian

$$H(\mathbf{q},\mathbf{p},\boldsymbol{m},\mathbf{Q}) = T(\mathbf{p},\boldsymbol{m}) + V(\mathbf{q},\mathbf{Q}), \tag{3.10}$$

where T refers to the total kinetic energy

$$T(\mathbf{p},\boldsymbol{m}) = \sum_i (T_i^{\mathrm{trans}}(\mathbf{p}_i) + T_i^{\mathrm{rot}}(\boldsymbol{m}_i)),$$

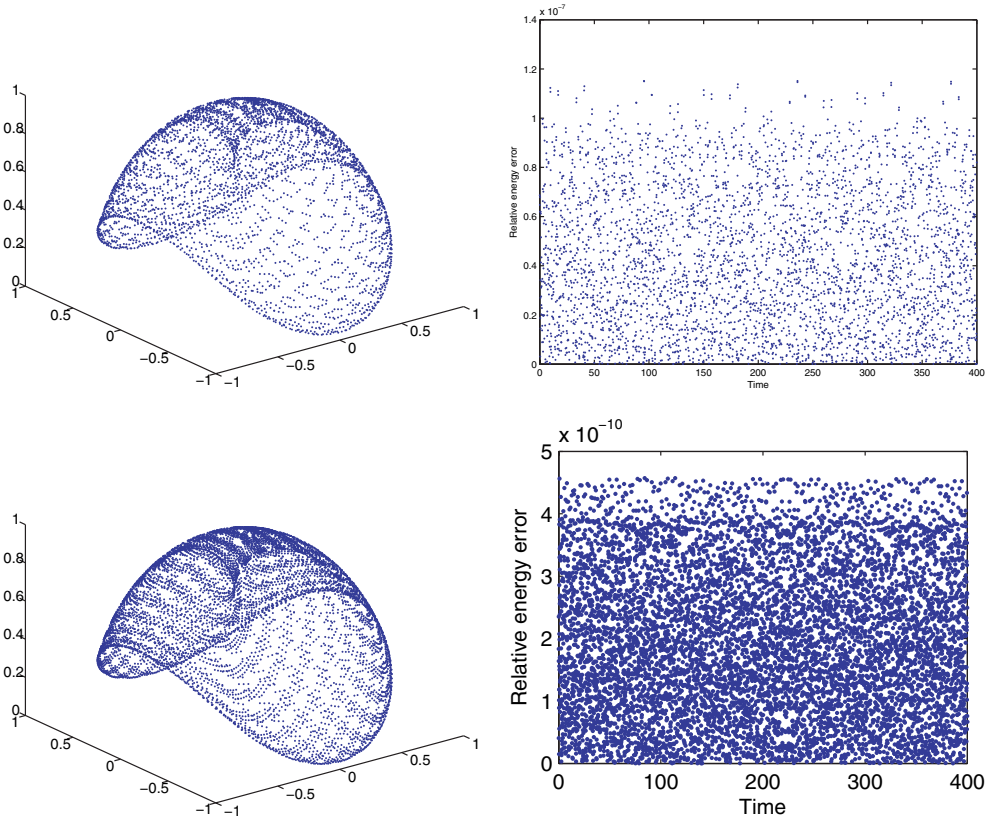

FIG. 3.6. *Satellite simulation. Left column: Center of mass* $(Q^T\mathbf{e}_3)$ *by the splitting method* $RKN6_{14}^a$ *with step size* $h = 0.1$ *(top) and* $h = 0.05$ *(bottom). Right: Relative error on the energy corresponding to the same step sizes. See text for details.*

consisting of the sum of the translational and rotational kinetic energies of each body

$$T_i^{\text{trans}}(\mathbf{p}_i) = \frac{\|\mathbf{p}_i\|^2}{2}, \qquad T_i^{\text{rot}}(\boldsymbol{m}_i) = \frac{1}{2}\boldsymbol{m}_i \cdot (I_i^{-1}\boldsymbol{m}_i),$$

where $I_i = \text{diag}(I_{i,1}, I_{,2}, I_{i,3})$ is the inertia tensor of the ith body, while V is the potential energy, describing the interaction between dipoles, that is assumed to depend on the position and orientation only. Furthermore, $V = \sum_{j>i} V_{i,j}$, where $V_{i,j}$ describes the interaction between dipole i and dipole j. We suppose that

$$V_{i,j}(\mathbf{q}_i, \mathbf{Q}_i, \mathbf{q}_j, \mathbf{Q}_j) = V_{i,j}^{\text{short}} + V_{i,j}^{\text{dip}},$$

where

$$V_{i,j}^{\text{short}} = 4\epsilon\left(\frac{\sigma}{r_{i,j}}\right)^{12}, \qquad \mathbf{r}_{i,j} = \mathbf{q}_i - \mathbf{q}_j, \quad r_{i,j} = \|\mathbf{r}_{i,j}\|,$$

describes the short range interaction between particles i and j, while

$$V_{i,j}^{\text{dip}} = \frac{1}{r_{i,j}^3}\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j - \frac{3}{r_{i,j}^5}(\boldsymbol{\mu}_i \cdot \mathbf{r}_{i,j})(\boldsymbol{\mu}_j \cdot \mathbf{r}_{i,j})$$

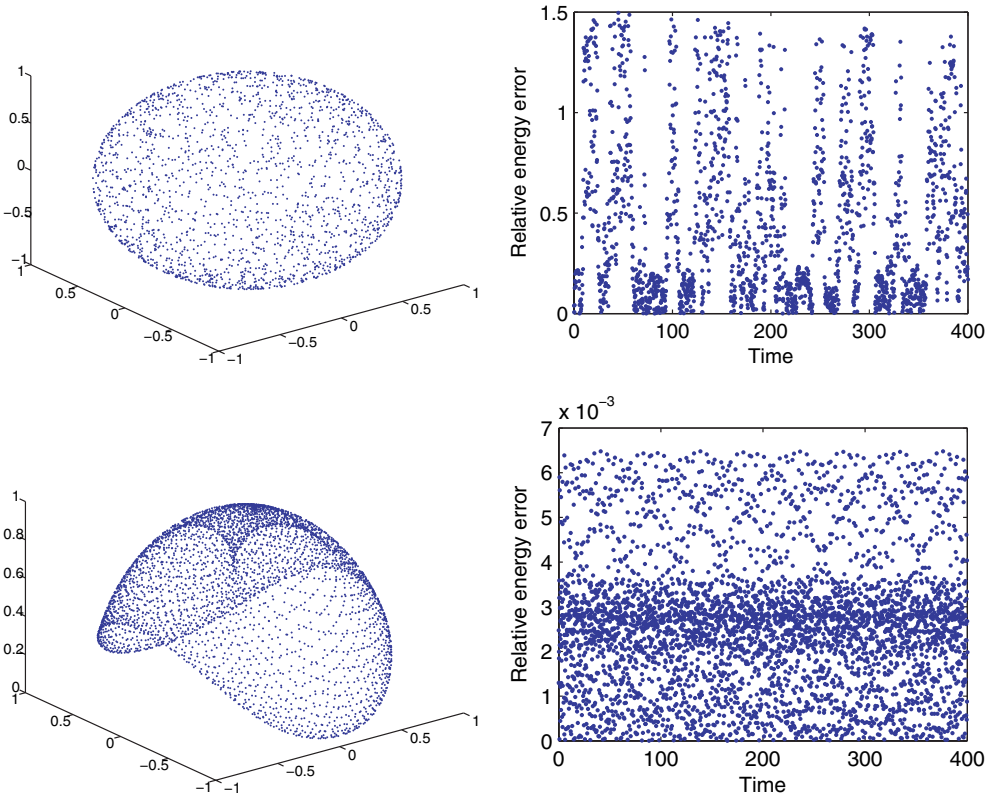

FIG. 3.7. *Satellite simulation. Left column: Center of mass* ($Q^T\mathbf{e}_3$) *by the splitting method MR with step size* $h = 0.1$ *(top) and* $h = 0.05$ *(bottom). Right: Relative error on the energy corresponding to the same step sizes. See text for details.*

is the term modeling the dipole interaction, where $\boldsymbol{\mu}_i$ is the orientation of the ith dipole vector. If $\bar{\boldsymbol{\mu}}_i$ is an initial fixed reference orientation for the dipole, then $\boldsymbol{\mu}_i = \mathbf{Q}_i\bar{\boldsymbol{\mu}}_i$.

The Hamiltonian (3.10) is separable, as the potential energy is independent of momenta and angular momenta. As before, we split the system as $H = T + V$, yielding

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m_i}, \\ \dot{\mathbf{p}}_i &= 0, \\ \dot{\boldsymbol{m}}_i &= \boldsymbol{m}_i \times (I_i^{-1}\boldsymbol{m}_i), \\ \dot{\mathbf{Q}}_i &= \mathbf{Q}_i\widehat{(I_i^{-1}\boldsymbol{m}_i)} \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} \dot{\mathbf{q}}_i &= 0, \\ \dot{\mathbf{p}}_i &= -\frac{\partial V}{\partial \mathbf{q}_i}, \\ \dot{\boldsymbol{m}}_i &= -\mathrm{rot}\left(\mathbf{Q}_i^\top \frac{\partial V}{\partial \mathbf{Q}_i}\right), \\ \dot{\mathbf{Q}}_i &= 0. \end{aligned} \tag{3.12}$$

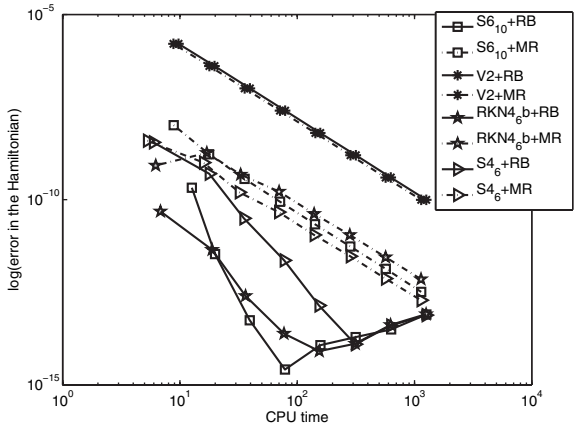

FIG. 3.8. *Error in the Hamiltoninan versus computational time for* 100 *particles. Several splitting methods are compared. See text for details.*

We approximate the original system with full Hamiltonian (3.10) by a composition of the flows of (3.11) and (3.12), by using some of the optimized splitting schemes introduced earlier. In [30] the authors use a similar approach. The main difference is in the choice of the splitting schemes (Störmer–Verlet and a fourth order Forest–Ruth-like scheme) and the implementation of the RB method. One of the standard methods, used in several packages for molecular dynamics simulations, for instance, the ORIENT package [27], is that described in [7]. The method consists of a Störmer–Verlet splitting plus a further splitting of the rigid body kinetic energy (that is, the MR method described earlier in section 3.3). Here we will denote the same method by V2+MR.

It is important to stress that, for a sufficiently large number of particles, approximating the rigid body equations by a inexpensive method, such as MR, or a more expensive one, such as the exact RB, is irrelevant, as the cost of this part grows only linearly with the number of particles. The computationally most demanding part in this simulation is the solution of (3.12), namely, the computation of the potential, whose cost grows quadratically with the number of particles.

This appears clearly in our first example: we compare different splitting methods for a system of 100 particles, for a relatively short time integration ($T_{\text{fin}} = 1$). All of the methods use a fixed step size, appropriately scaled for each splitting scheme, to require the same number of function evaluations. For the reference method, the V2+MR, we use step size $h = 10^{-1} \times 1/2^i$ for $i = 0, \ldots, 7$; i.e., for the largest step size $h = 0.1$, one has 10 potential evaluations, and thus the x-axis in Figure 3.8 can be interpreted as a number of function evaluations as well. Similarly, the sixth order splitting method S6$_{10}$+RB, with 10 internal stages requiring potential evaluations, is implemented with step size $h = 1$. The results of the simulation are displayed in Figure 3.8. The methods are implemented by using the RB method (solid line) and the MR method (dashed-dotted line). Coalescence of stages is exploited for all methods.

The initial conditions for the experiment were taken as follows: the masses m_i are chosen to be 1, $\mathbf{q}_i = N \times \texttt{randn}(3,1)$, $N = 100$ being the number of particles and $\texttt{randn}(3,1)$ a vector with random components (Gaussian distribution) between -1 and 1; $\mathbf{p}_i = 0, \boldsymbol{m}_i = 0$, $\mathbf{Q_i}$ random orthogonal matrix, $\boldsymbol{\mu}_i = (0,1,1)^T$, $\sigma = \epsilon = 1$,

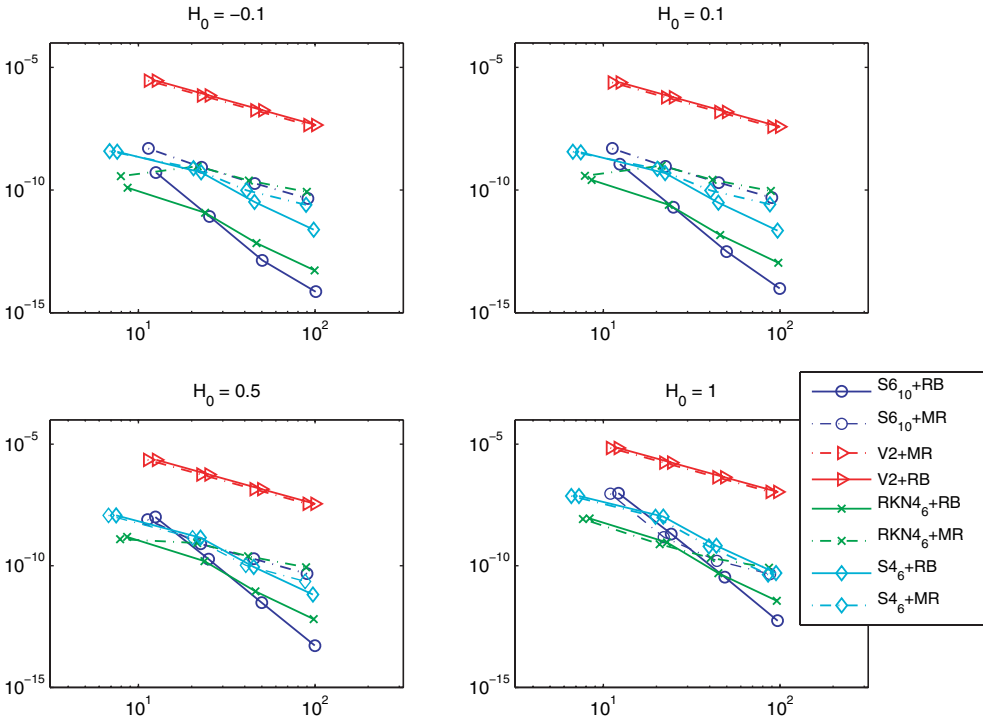

FIG. 3.9. *Average errors versus CPU time for different values of the energy H_0 and 100 runs per each of the step sizes $1, 1/2, 1/4$, and $1/8$. Number of particles $N = 125$. For small energy values, the splitting methods based on the exact RB integrator perform better than those with the MR splitting. For higher values of the energy, the error due to the splitting $H = T + V$ is much higher than the error for the RB part, and it dominates the total error.*

with a resulting energy $H_0 = 0.14134185611814$. The moments of inertia are those of water ($I_1 = 1, I_2 = 1.88$, and $I_3 = 2.88$).

In the next numerical example (Figure 3.9), we test the same methods for different energies. The initial conditions are chosen as follows: we take 125 particles that we position on a lattice of dimension $5 \times 5 \times 5$. The initial positions are then perturbed by 1% (Gaussian normal distribution). The initial orientations are random orthogonal matrices. With these parameters, we compute the initial energy, and then we change the linear momentum of the particles in positions $\mathbf{q}_1 = (1, 1, 1)^T$ and $\mathbf{q}_{125} = (5, 5, 5)^T$ to achieve the target energy H_0. For each step size $h = 1, 1/2, 1/4, 1/8$ of the basic method $\mathrm{SR6}_{10}$, we perform 100 simulations (choosing every time a different initial condition), and we average the error and the computational time (arithmetic mean).

Finally, having observed that Nyström schemes behave very well for this class of problems, the method $\mathrm{RKN4}_6^b$ is compared to $\mathrm{RKN6}_{14}^a$ in Figure 3.10. The number of function evaluations for the two methods is the same. The initial conditions are as before, except for the number of averages (which is 1) and the time of integration, with $T_{\mathrm{fin}} = 10$.

These experiments indicate that the main source of error for this problem is the splitting $H = T + V$. This seems consistent with conclusions on the water simulations in [30]. In particular, a lower global error as a result of using the exact RB integrator becomes visible only in the step size asymptotic regime (as $h \to 0$; see Figure 3.8). Furthermore, how small the step size h must be, to see the positive effect of the exact

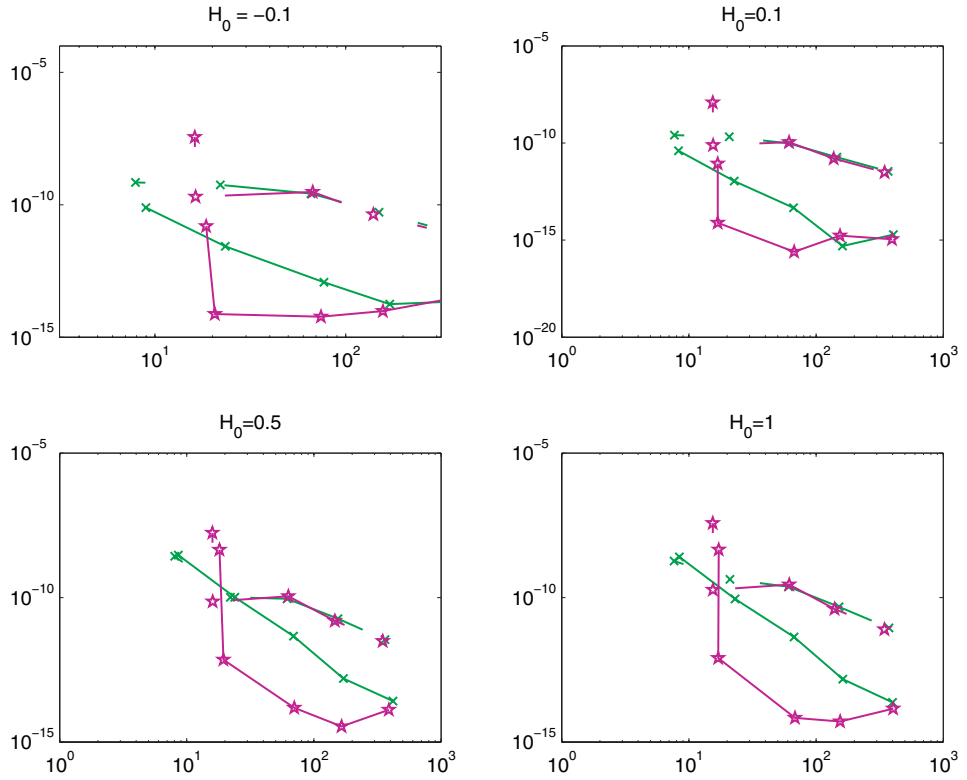

FIG. 3.10. *Error versus CPU time. Comparison of two RKN splittings of order* 4 *and* 6*, on the interval* $[0, 10]$*,* 125 *particles, for some initial conditions. The sharp increase of the error for the sixth order method is due to the fact that the step size is too large. The* $RKN4_6$ *method (crosses with solid and dashed-dotted lines) is the same as in Figure* 3.9.

RB integrator, seems to depend on the total energy of the system. For low energies, the error with the exact method becomes smaller at a larger step size. For higher energies, the error with the exact method becomes smaller only at very small values of h; see Figures 3.9–3.10.

Our conclusion is that the use of an exact RB integrator is favorable for simulations where higher precision is required (for instance, low energy). For higher energies, the effect of having an exact integrator for the RB part appears to be less relevant unless other techniques are used.

4. Conclusions. The main purpose of this paper has been to understand whether and when the use of an exact solution of the free rigid body equations as a component of splitting methods is a competitive geometric integrator.

We have reviewed various algorithms for the computation of the exact solution of the free rigid body equations providing a common framework. We have implemented two concrete approaches based on rotation matrices and quaternions. The algorithms require the computation of an elliptic integral of the third kind, which we either compute to machine accuracy (exact methods) or approximate by Gauss–Legendre quadrature (semiexact methods). We have performed numerous experiments comparing these methods to preprocessed discrete Moser–Veselov methods (experiments on the free rigid body problem) and to various splitting methods (experiments on torqued rigid bodies).

In conclusion, the exact methods, though more expensive, are very robust and behave uniformly well for all choices of the principal moments of inertia and initial conditions, independently of the step size of integration.

If cost is an issue, *semiexact* methods are a good compromise. They are much cheaper than the exact ones, while sharing most of the geometric properties[2] and being robust for large step sizes and arbitrary values of the principal moments of inertia. This is an advantage compared to implicit methods using fixed-point iteration that might require small step sizes to converge (e.g., the cheaper DMV).

Our conclusion is that the implementation of the exact solution of the free rigid body is competitive as a numerical approach and has the advantage that it can be used as a building block for splitting methods of high order.

When used as a component of splitting methods the exact and semiexact methods are definitively of interest in the case of perturbed free rigid body problems. In molecular dynamics simulations, using the exact rigid body motion gives a clear advantage compared to other splittings only in the low energy case. However, even in the general case, the cost due to any rigid body integrator (approximate or exact) is growing only linearly as a function of the number of particles, while the number of interparticle force evaluations is growing quadratically and dominates the overall computational cost.

Appendices.

Appendix A. Jacobi elliptic functions. We collect here a few facts about the elliptic functions we use in the article. Given $0 \le k < 1$, the function

$$\varphi \mapsto F(\varphi, k) := \int_0^{\varphi} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}} \tag{4.1}$$

is called an (incomplete) *elliptic integral of the first type* with modulus k and is a diffeomorphism $\mathbb{R} \to \mathbb{R}$. Its inverse $F(\,\cdot\,, k)$ is an odd function

$$\operatorname{am}(\,\cdot\,, k) : \mathbb{R} \to \Big(-\frac{\pi}{2}, \frac{\pi}{2} \Big),$$

which is called the *amplitude* of modulus k. The *Jacobi elliptic functions* sn and cn of modulus k are the functions $\mathbb{R} \to [-1, 1]$ defined as

$$\operatorname{sn}(u, k) = \sin(\operatorname{am}(u, k)), \qquad \operatorname{cn}(u, k) = \cos(\operatorname{am}(u, k)),$$

respectively, and are periodic of period $4K(k)$, where $K(k) = F(\frac{\pi}{2}, k)$ (the so-called complete elliptic integral of the first type of modulus k). Moreover,

$$\operatorname{dn}(u, k) = \sqrt{1 - k^2 \operatorname{sn}(u, k)^2}, \qquad \operatorname{sd}(u, k) = \frac{\operatorname{sn}(u, k)}{\operatorname{dn}(u, k)}.$$

For given k, the u-derivatives of these functions satisfy $\operatorname{sn}' = \operatorname{cn} \operatorname{dn}$, $\operatorname{cn}' = -\operatorname{sn} \operatorname{dn}$, and $\operatorname{dn}' = -k^2 \operatorname{sn} \operatorname{cn}$.

The (incomplete) *elliptic integral of the third kind* with modulus $0 < k \le 1$ and parameter $n \in \mathbb{R}$ is the function $\Pi(\,\cdot\,, n, k) : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ defined by

$$\Pi(\varphi, n, k) := \int_0^{\varphi} \frac{d\theta}{(1 - n \sin^2 \theta) \sqrt{1 - k^2 \sin^2 \theta}} \tag{4.2}$$

[2]Symplecticity is lost, but the methods are time-reversible as long as the underlying quadrature is symmetric. Also, the DMV methods are time-reversible but not symplectic.

(Legendre form) or, equivalently,

$$\Pi(\varphi, n, k) = \int_0^{F(\varphi,k)} \frac{ds}{1 - n\,\mathrm{sn}(s,k)^2}\,.$$

Appendix B. Coefficients of the Gauss quadrature. For completeness, we report the coefficients of the Gaussian quadrature of order 10 shifted to the interval $[0,1]$:

$$\begin{aligned} a_1 &= 0.04691007703067, & b_1 &= 0.11846344252809,\\ a_2 &= 0.23076534494716, & b_2 &= 0.23931433524968,\\ a_3 &= 0.5, & b_3 &= 0.28444444444444,\\ a_4 &= 0.76923465505284, & b_4 &= b_2,\\ a_5 &= 0.95308992296933, & b_5 &= b_1. \end{aligned} \tag{4.3}$$

For the qudrature of order 6 and 8 the coefficients have closed form and can be found, for instance, in [1].

Appendix C. Coefficients of the splitting schemes. Given the differential equation

$$y' = F(y) = A(y) + B(y),$$

denote by $\varphi_\tau^{[F]}$ the flow of the vector-field F from time t to time $t+\tau$. Given a numerical approximation $y^{(j)} \approx y(t_j)$, we consider symmetric splitting schemes of the type

$$y^{(j+1)} = \varphi_{a_1 h}^{[A]} \circ \varphi_{b_1 h}^{[B]} \circ \varphi_{a_2 h}^{[A]} \circ \cdots \circ \varphi_{a_{m+1} h}^{[A]} \circ \cdots \varphi_{b_1 h}^{[B]} \circ \varphi_{a_1 h}^{[A]}\; y^{(j)},$$

where $h = t_{j+1} - t_j$. A typical splitting is obtained by separating the contributions arising from the kinetic (A) and potential (B) energy of the system. For this reason, (twice) the number s of the coefficients b_i is called the *stage number* of the splitting method. The effective error is defined as $E_f = s\sqrt[p]{||\mathbf{c}||_2}$, where $\mathbf{c}$ is the vector of coefficients of the elementary differentials of the leading error term and p is the order of the method. We refer to [4, 21] for background and notation.

For completeness, we report the coefficients of the methods used in this paper.

Störmer–Verlet scheme (V2):

$$a_1 = 1/2, \qquad b_1 = 1, \tag{4.4}$$

(order 2, one stage).

$\mathrm{S6}_{10}$ method (order 6, 10 stages, effective error $E_f = 1.12$):

$$\begin{aligned} a_1 &= 0.0502627644003922, & b_1 &= 0.148816447901042,\\ a_2 &= 0.413514300428344, & b_2 &= -0.132385865767784,\\ a_3 &= 0.0450798897943977, & b_3 &= 0.067307604692185,\\ a_4 &= -0.188054853819569, & b_4 &= 0.432666402578175,\\ a_5 &= 0.541960678450780, & b_5 &= 1/2 - (b_1 + \cdots + b_4),\\ a_6 &= 1 - 2(a_1 + \cdots + a_5). \end{aligned} \tag{4.5}$$

$\mathrm{S4}_6$ (order 4, 6 stages, effective error $E_f = 0.56$):

$$\begin{aligned} a_1 &= 0.07920369643119565, & b_1 &= 0.209151106613362,\\ a_2 &= 0.353172906049774, & b_2 &= 0.143851773179818,\\ a_3 &= -0.04206508035771952, & b_3 &= 1/2 - (b_1 + b_2),\\ a_4 &= 1 - 2(a_1 + a_2 + a_3). \end{aligned} \tag{4.6}$$

The splittings above are generic in the sense that the A and B part are interchangeable. This is not the case for the next methods, which are based on Nyström schemes for separable Hamiltonians.

RKN4$_6^b$ (order 4, (7)6 stages, effective error $E_f = 0.28$):

$$\begin{aligned} b_1 &= 0.0829844064174052, & a_1 &= 0.245298957184271, \\ b_2 &= 0.396309801498368, & a_2 &= 0.604872665711080, \\ b_3 &= -0.0390563049223486, & a_3 &= 1/2 - (a_1 + a_2), \\ b_4 &= 1 - 2(b_1 + b_2 + b_3). \end{aligned} \tag{4.7}$$

RKN6$_{14}^a$ (order 6, 14 stages, effective error $E_f = 0.63$):

$$\begin{aligned} a_1 &= 0.0378593198406116, & b_1 &= 0.09171915262446165, \\ a_2 &= 0.102635633102435, & b_2 &= 0.183983170005006, \\ a_3 &= -0.0258678882665587, & b_3 &= -0.05653436583288827, \\ a_4 &= 0.314241403071477, & b_4 &= 0.004914688774712854, \\ a_5 &= -0.130144459517415, & b_5 &= 0.143761127168358, \\ a_6 &= 0.106417700369543, & b_6 &= 0.328567693746804, \\ a_7 &= -0.00879424312851058, & b_7 &= 1/2 - (b_1 + \cdots + b_6), \\ a_8 &= 1 - 2(a_1 + \cdots + a_7). \end{aligned} \tag{4.8}$$

Acknowledgments. The authors thank E. Hairer, B. Carlson, A. Giacobbe, and E. Karatsuba for useful discussions and comments. We acknowledge the kind hospitality and support of the Newton Institute of Mathematical Sciences in Cambridge, UK. Special thanks to A. Iserles.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Appl. Math. Ser. 55 (reprint of the 1972 edition) Dover, New York, 1992.

[2] P. E. Appell, *Traité de Mécanique Rationnelle*, Vol. 2, Gauthier–Villars, Paris, 1924/26.

[3] G. Benettin, A. M. Cherubini, and F. Fassò, *A changing-chart symplectic algorithm for rigid bodies and other Hamiltonian systems on manifolds*, SIAM J. Sci. Comput., 23 (2001), pp. 1189–1203.

[4] S. Blanes and P. C. Moan, *Practical symplectic partitioned Runge–Kutta and Runge–Kutta–Nyström methods*, J. Comput. Appl. Math., 142 (2002), pp. 313–330.

[5] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists*, Grundlehren Math. Wiss. 67, 2nd ed., Springer-Verlag, New York, 1971.

[6] E. Celledoni and N. Säfström, *Efficient time-symmetric simulation of torqued rigid bodies using Jacobi elliptic functions*, J. Phys. A, 39 (2006), pp. 5463–5478.

[7] A. Dullweber, B. Leimkuhler, and R. McLachlan, *Symplectic splitting methods for rigid body molecular dynamics*, J. Chem. Phys., 107 (1997), pp. 5840–5851.

[8] K. Efstathiou and D. Sadovskii, *No Polar Coordinates. Based on Lectures by R. Cushman*, London Math. Soc. Lecture Note Ser. 306, Geometric Mechanics and Symmetry, Cambridge University Press, Cambridge, 2005.

[9] F. Fassò, *Comparison of splitting algorithm for the rigid body*, J. Comput. Phys., 189 (2003), pp. 527–538.

[10] F. Fassò, *Superintegrable Hamiltonian systems: Geometry and perturbations*, Acta Appl. Math., 87 (2005), pp. 93–121.

[11] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer Ser. Comput. Math. 31, Springer-Verlag, New York, 2002.

[12] E. Hairer and G. Vilmart, *Preprocessed discrete Moser-Veselov algorithm for the full dynamics of a rigid body*, J. Phys. A, 39 (2006), pp. 13225–13235.

[13] C. G. J. Jacobi, *Sur la rotation d'un corps*, J. Reine Angew. Math., 39 (1949), pp. 293–350.

[14] S. H. Morton, Jr., J. L. Junkins, and J. N. Blanton, *Analytical solutions for Euler parameters*, Celestial Mech., 10 (1974), pp. 287–301.

[15] I. I. Kosenko, *Integration of the equations of a rotational motion of rigid body in quaternion algebra. The Euler case*, J. Appl. Math. Mech., 62 (1998), pp. 193–200.
[16] D. F. Lawden, *Elliptic Functions and Applications*, Appl. Math. Sci. 80, Springer-Verlag, New York, 1989.
[17] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge Monogr. Appl. Comput. Math. 14, 1st ed., Cambridge University Press, London, 2004.
[18] M. Leok, T. Lee, and N. H. McClamroch, *Attitude maneuvers of a rigid spacecraft in a circular orbit*, in Proceedings of the IEEE American Control Conference, 2006, pp. 1742–1747.
[19] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry, texts a basic exposition of classical mechanical systems*, in Appl. Math. 17, 2nd ed., Springer-Verlag, New York, 1999.
[20] R. I. McLachlan, *Explicit Lie-Poisson integration and the Euler equations*, Phys. Rev. Lett., 71 (1993), pp. 3043–3046.
[21] R. I. McLachlan and G. R. W. Quispel, *Splitting methods*, Acta Numer., 11 (2002), pp. 341–434.
[22] R. I. McLachlan and A. Zanna, *The discrete Moser–Veselov algorithm for the free rigid body, revisited*, Found. Comput. Math., 5 (2005), pp. 87–123.
[23] J. Wm. Mitchell, *A Simplified Variation of Parameters Solution for the Motion of an Arbirarily Torqued Mass Asymmetric Rigid Body*, Ph.D. thesis, University of Cincinnati, 2000.
[24] J. Moser and A. Veselov, *Discrete versions of some classical integrable systems and factorization of matrix polynomials*, J. Comm. Math. Phys., 139 (1991), pp. 217–243.
[25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran* 77 *and Fortran* 90*, The Art of Scientific and Parallel Computing*, 2nd ed., Cambridge University Press, Cambridge, 1996.
[26] S. Reich, *Symplectic integrators for systems of rigid bodies. Integration algorithms and classical mechanics*, Fields Inst. Commun., 10 (1996), pp. 181–191.
[27] A. J. Stone, A. Dullweber, M. P. Hodges, P. L. A. Popelier, and D. J. Wales, *ORIENT Version* 3.2: *A Program for Studying Interaction between Molecules*, Cambridge University, London, 1995.
[28] J. Touma and J. Wisdom, *Lie–poisson integrators for rigid body dynamics in the solar system*, Astr. J., 107 (1994), pp. 1189–1202.
[29] R. van Zon and J. Schofield, *Numerical implementation of the exact dynamics of free rigid bodies*, J. Comput. Phys., 225 (2007), pp. 145–164.
[30] R. van Zon and J. Schofield, *Symplectic algorithms for simulations of rigid body systems using the exact solution of free motion*, Phys. Rev. E, 75 (2007), 056701.
[31] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, London, 1937.
[32] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990). pp. 262–268.