PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

Article submitted to journal

Subject Areas:

xxxxx, xxxxx, xxxx

Keywords:

N-Body Problem; Point-vortex dynamics; Hamiltonian Systems; Newton Laws; Surfaces of Revolution; Equivalence Principle

Author for correspondence:

Stefanella Boatto

e-mail: boatto.stefanella@gmail.com

THE ROYAL SOCIETY PUBLISHING

N-body Dynamics on Closed Surfaces; The Axioms of Mechanics

Stefanella Boatto[1], David G. Dritschel[2] and Rodrigo G. Schaefer[3]

[1]Departamento de Matemática Aplicada, Instituto de Matemática, Universidade Federal de Rio de Janeiro, Rio de Janeiro, RJ, CEP 21941-909, Brazil

[2]School of Mathematics and Statistics, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Scotland, UK

[3]Department de Matemàtica Aplicada, Facultat de Matemàtiques i Estadística, Universitat Politècnica de Catalunya Pau Gargallo, 5. Barcelona 08028,Spain

A major challenge for our understanding of the mathematical basis of particle dynamics is the formulation of N-body and N-vortex dynamics on Riemann surfaces. In this paper we show how the two problems are, in fact, closely related when considering the role played by the intrinsic geometry of the surface. This enables a straightforward deduction of the dynamics of point masses, using recently-derived results for point vortices on general closed differentiable surfaces M endowed with a metric g. We find, generally, that Kepler's laws do not hold. What is more, even Newton's first law (the Law of Inertia) fails on closed surfaces having variable curvature (e.g. the ellipsoid).

1. Introduction

Every theoretical model relies on what mathematicians call *axioms* and physicists call *working hypotheses*, the foundations of theory. In geometry, this is what characterizes for example Euclidean geometry, where parallel geodesics — straight lines — never cross, and spherical geometry, where parallel geodesics — great circles — always cross. In Newtonian mechanics in Euclidean spaces, a so called *Mechanical System* is one that verifies the three Newtonian laws, which we could consider as the axioms of mechanics [1].

In this article we show that Newton's first Law, the *Law of inertia*, is not universal. In particular, on compact surfaces without boundaries that are conformal to the sphere, we show that the Law of Inertia is only valid in special geometries, namely surfaces with constant Gaussian curvature. We can therefore view Newton's first law as playing a similar role for a Mechanical System, on a sphere or on a surface conformal to the sphere, as Euclid's fifth postulate distinguishing planar geometry from spherical geometry.

The crucial insight is to regard a surface in its intrinsic geometry, *not* as one embedded in $\mathbb{R}^3$, as done for example in [10]. This leads to significant differences in the formulation of the dynamics, as is already known in the case of point vortices [13]. Here, particular attention is given to closed surfaces of revolution, for example the ellipsoid of revolution and the bean surface, shown in Figure 1. Such surfaces permit a straightforward, explicit formulation [13].

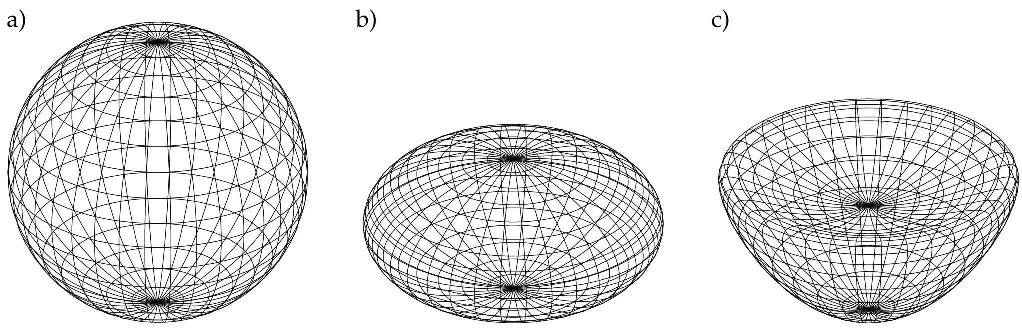

Figure 1. a) a sphere, b) an oblate ellipsoid of revolution and c) a bean-shape surface with parameters $a = 0.6$ and $b = 0.4$ (see [13]).

2. The first axiom of Newtonian mechanics revisited

In 1687, Isaac Newton published in his Principia [21] three famous laws which have become widely regarded as the "axioms of mechanics". The first law, the Law of Inertia, states:

The vis insita, or innate force of matter, is a power of resisting by which every body, as much as in it lies, endeavours to preserve its present state, whether it be of rest or of moving uniformly forward in a straight line.

On a general surface, the analogue of a straight line is a *geodesic* [11]. It is well known that a particle on a plane either remains at rest, if its initial velocity is zero, or continues in a straight line with uniform velocity equal to its initial velocity. On a sphere we have a similar scenario as on the plane. Either a particle remains at rest or it travels at constant speed around a great circle [6,24]. Both the plane and the sphere are surfaces with constant *Gaussian curvature* [11]. What happens if for example we deform the sphere into an ellipsoid of revolution? In this case, as shown in Figure 2 and as explained in §8, a particle initially at rest generally begins to move along a meridian! It then stops and reverses its direction of motion. The particle oscillates about the equator on an oblate ellipsoid (shown), and about the closest pole on a prolate ellipsoid.

This oscillatory motion can be understood as follows. The particle is initially at a co-latitude θ_o, which divides the surface into two, generally-unequal regions. Since gravity is a central force, the uniform background mass (which is required on a closed surface as discussed below), acts as if it were concentrated at two points along the axis of revolution between the poles. The net force is generally unbalanced and sets the particle in motion. On an oblate ellipsoid, the particle moves towards the equator where the net force changes sign. The momentum of the particle then carries it to $\theta = \pi - \theta_o$ before it reverses direction and returns to its initial position. On a prolate ellipsoid, instead the particle oscillates about the nearest pole.

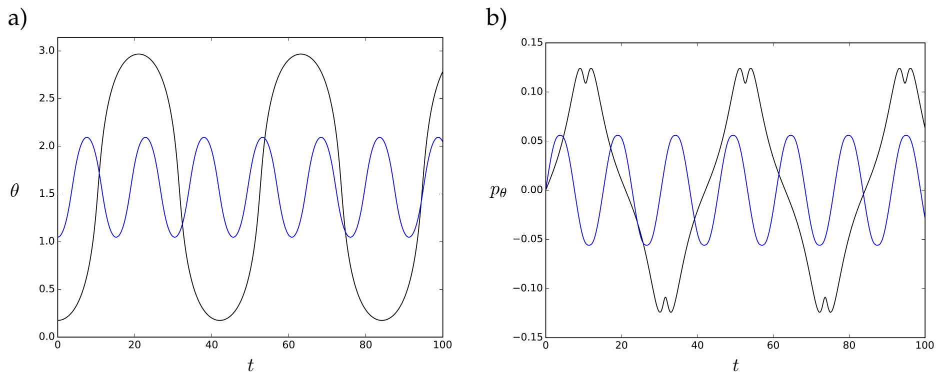

Figure 2. Dynamics of a single particle, initially at rest, on an oblate ellipsoid of revolution. a) Time evolution of the co-latitude θ for two cases starting at $\theta = \theta_o = 10°$ (black curve), and at $\theta = \theta_o = 60°$ (blue curve). b) Corresponding time evolution of the meridional momentum, p_θ.

3. Surfaces of revolution

As in [13] we consider a surface of revolution M (about the vertical z axis) which is a deformation of a sphere of radius R. The Cartesian coordinates (x, y, z) in $\mathbb{R}^3$ of any point on M may be expressed as functions of two surface coordinates θ and ϕ, co-latitude and longitude, respectively. For surfaces of revolution it is sufficient to take

$$x = \rho(\theta)\,\cos\varphi; \qquad y = \rho(\theta)\,\sin\varphi; \quad z = \zeta(\theta),$$

where $\rho(\theta)$ and $\xi(\theta)$ are specified functions of θ – which in the plane yz describe the curve generating the surface. Without loss of generality, we may take $0 \le \varphi \le 2\pi$ and $0 \le \theta \le \pi$ over S. Note that:

- for the sphere:
$$\rho(\theta) = R\,\sin\theta, \qquad \zeta(\theta) = R\,\cos\theta; \tag{3.1}$$
- for the ellipsoid of revolution:
$$\rho(\theta) = R\,\sin\theta, \qquad \zeta(\theta) = b\,R\,\cos\theta, \tag{3.2}$$
where b is the height-to-width aspect ratio; and
- for the bean surface:
$$\rho(\theta) = R\,\sin\theta, \qquad \zeta(\theta) = R\,(a\,\sin^2\theta + b\,\cos\theta), \tag{3.3}$$
where a is an asymmetry parameter.

(a) Metric and metric tensor of a surface of revolution

The differential distance ds – also called the *metric* of M – between two points on S is

$$\mathrm{d}s^2 = |\mathrm{d}x|^2 = \mathrm{d}x^2 + dy^2 + dz^2 = [(\rho')^2 + (\zeta')^2]\,\mathrm{d}\theta^2 + \rho^2\,\mathrm{d}\varphi^2, \tag{3.4}$$

where primes denote differentiation with respect to θ. From Eqs. (7.5)-(3.3) it follows that

- for the sphere of radius R:
$$\mathrm{d}s^2 = R^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2); \tag{3.5}$$
- for the ellipsoid of revolution:
$$\mathrm{d}s^2 = R^2\left\{\left(\cos^2\theta + b^2\sin^2\theta\right)\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right\}; \tag{3.6}$$

- for the bean surface:

$$ds^2 = R^2\left\{\left[\cos^2\theta + \sin^2\theta\,(2a\,\cos\theta - b)^2\right]\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right\}. \tag{3.7}$$

4. The motion of a test particle in a gravitational field

By analogy with the fluid dynamics problem [13], we start by considering the motion of a test particle on a surface M with a metric g. We suppose that the density of matter $\rho(\boldsymbol{r},t)$ is given. To put things into context, we may think of a satellite of mass m_0 in the gravitational field of the planets in our solar system, with masses $m_1, ..., m_N$. We then assume $m_0 \ll m_j,\ j = 1, ..., N$. This is what Poincaré called the *restricted N-body problem* [22]. Following Poincaré, we assume that the presence of the satellite does not affect the motion of the planets. Moreover, for the moment, we also assume that the planets' trajectories, $\boldsymbol{r}_1(t)$, ..., $\boldsymbol{r}_N(t)$, are known. The question then is: how do we deduce the motion of the satellite on a general surface M?

By analogy with Maxwell's laws (see Jackson [18]), we start with the fundamental equations of a mechanical system in the presence of a central force. Let $\boldsymbol{a}(\boldsymbol{r},t)$ be the acceleration field surrounding the satellite located at $\boldsymbol{r}$ at time t. Since the force (per unit mass) is irrotational (curl free) and attractive, $\boldsymbol{a}$ must satisfy essentially the same laws satisfied by an electric field $\boldsymbol{E}$, namely:

$$\begin{aligned} &i) && curl(\boldsymbol{a}) = 0, \\ &ii) && div(\boldsymbol{a}) = -\frac{\rho(\boldsymbol{r},t)}{\epsilon}, \\ &iii) && \frac{\partial \boldsymbol{a}}{\partial t} + \frac{1}{\epsilon}\boldsymbol{J} = 0, \end{aligned} \tag{4.1}$$

where the minus sign in the second equation reflects the fact that gravity is a force of attraction [6]. Above, $\rho(\boldsymbol{r},t)$ is the mass density, ϵ is a suitable constant analogous to permittivity in electrostatics, and $\boldsymbol{J}$ is the mass current. The last equation is nothing more than the statement of mass conservation. The second equation can be re-expressed as

$$div(\boldsymbol{a}) = -\frac{1}{\epsilon}\,\rho(\boldsymbol{r},t) = -\gamma\,\rho(\boldsymbol{r},t). \tag{4.2}$$

where the gravitational constant γ satisfies $\gamma = 1/\epsilon$.

Considering surfaces that are smooth deformations of the sphere $\mathbb{S}^2$ (cf. Figure 1), Eq. (4.1) allows us to express the acceleration field as the gradient of an unknown function Φ:

$$\boldsymbol{a} = -grad(\Phi),$$

where Φ plays the role of the satellite gravitational potential. It is determined by substituting the above into Eq. (4.2), giving

$$\Delta_g \Phi = \gamma\rho(\boldsymbol{r},t) \tag{4.3}$$

where Δ_g is the Laplace-Beltrami operator, generalizing the Laplacian for a surface with metric g [19].

Remarks

1) Given the density of matter $\rho(\boldsymbol{r},t)$ – i.e. given the dynamics of the planets – and the metric g of the surface under consideration, the gravitational potential of the satellite, Φ, is the solution of the Poisson equation (4.3). A similar equation arises for the streamfunction Ψ of an incompressible fluid [13] for a given vorticity field $\omega(\boldsymbol{r},t)$, i.e. $\Delta_g\Psi = \omega$, see Appendix A.

2) The Poisson equation (4.3) is linear. Hence, there exist a fundamental solution $G(\boldsymbol{r},\boldsymbol{r}')$, called the Green function [17], in terms of which the solution of Eq. (4.3) is

$$\Phi(\boldsymbol{r},t)=\gamma\iint_M G(\boldsymbol{r},\boldsymbol{r}')\,\rho(\boldsymbol{r}',t)\,d\boldsymbol{r}'. \tag{4.4}$$

For $\mathbb{R}^3$, and for surfaces conformal to the plane, $G(\boldsymbol{r},\boldsymbol{r}')$ is the solution of

$$\Delta_g G(\boldsymbol{r},\boldsymbol{r}')=\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{4.5}$$

where $\delta(\boldsymbol{r}-\boldsymbol{r}')$ is the usual Dirac delta distribution. More specifically (see [4,15,19]), for $\mathbb{R}^3$, the *punctured* sphere $\mathbb{S}^2_p$, the plane $\mathbb{R}^2$, and the hyperbolic surface $\mathbb{H}^2$ (of constant negative Gaussian curvature K) we have that

$$\begin{array}{rcll}
G_{\mathbb{R}^3}(\boldsymbol{r},\boldsymbol{0}) &=& -\dfrac{1}{4\pi}\dfrac{1}{||\boldsymbol{r}||}, & \boldsymbol{r}=(x,y,z)\\
G_{\mathbb{S}^2_p}(\boldsymbol{r},\boldsymbol{0}) &=& \dfrac{1}{2\pi}\ln\left[\tan\left(\dfrac{\sqrt{K}}{2}r\right)\right], & \boldsymbol{r}=(\varphi,\theta),\quad r=\dfrac{\theta}{\sqrt{K}}\\
G_{\mathbb{R}^2}(\boldsymbol{r},\boldsymbol{0}) &=& \dfrac{1}{2\pi}\ln r, & \boldsymbol{r}=(x,y),\quad r=\sqrt{x^2+y^2}\\
G_{\mathbb{H}^2}(\boldsymbol{r},\boldsymbol{0}) &=& \dfrac{1}{2\pi}\ln\left[\tanh\left(\dfrac{\sqrt{|K|}}{2}r\right)\right], & \boldsymbol{r}=(\varphi,\theta),\quad r=\dfrac{\theta}{\sqrt{|K|}}
\end{array}$$

respectively, where r is the geodesic distance. Notably, for the punctured sphere [4], the Green function converges to the planar one in the limit $K\to 0$ (or, equivalently, the radius of the sphere $R=1/\sqrt{K}\to\infty$):

$$2\pi G_{\mathbb{S}^2_p}(\boldsymbol{r},0)=\ln\left[\tan\left(\sqrt{K}r/2\right)\right]+\ln(\sqrt{K}/2)=\ln r+\frac{1}{12}Kr^2+O(K^2r^4),$$

for small K and fixed r, such that $K\,r^2\ll 1$.[1]

Up to this point, we have considered surfaces with their *intrinsic* geometry and not as embedded in $\mathbb{R}^3$. We have addressed the question: given the surface metric and the distribution of matter, can we deduce the mass dynamics? However, it is also possible to consider the dynamics from an *extrinsic* geometry point of view, i.e. now regarding M as embedded in $\mathbb{R}^3$. One way to do so is to restrict the $\mathbb{R}^3$ potential to the plane $\mathbb{R}^2$ and pull it back to the sphere, by the inverse of a stereographic projection from the South Pole ($r=\tan(\theta/2)$). In this case, the Green function for the unit punctured sphere is

$$\tilde{G}_{\mathbb{S}^2_p}(\boldsymbol{r})=-\frac{1}{4\pi}\frac{1}{||\boldsymbol{r}||}=-\frac{1}{4\pi}\,\cot(\theta/2)\,.$$

Such a potential is a potential for the punctured sphere, but it is not the solution of the Poisson equation (4.5), i.e. it is *not* the potential associated with the intrinsic geometry of the surface.

Moreover, various authors in the literature (cf. Koslov *et al.* (1991) [20], Borisov *et al.* (2004) [8] and Diacu (2012) [10]) use

$$\tilde{G}_{\mathbb{S}^2}(\boldsymbol{r})=-k\,\cot\theta$$

for the Green function of a sphere, where k is a constant. In this case, observe that if we hold one mass fixed at the north pole and consider a second mass at a co-latitude θ, then the force on the second mass is

$$\tilde{F}(\theta)=\tilde{\gamma}\,\frac{m_1m_2}{\sin^2\theta}\,,$$

which exhibits a minimum at the equator but increases without bound toward each pole (see Figure 3). This does not enable one to consider the equilibrium configuration consisting of one

[1] As the Green function is defined up to a constant, see Appendix A, we have taken the liberty to add $\ln(\sqrt{K}/(4\pi))$ to it.

mass at the north pole and a second mass at the south pole, a configuration for which we recover with our potential (see remark 3 below) the corresponding force

$$F(\theta) = \frac{\gamma}{4\pi}\,\frac{m_1 m_2 \sin\theta}{1-\cos\theta}\,.$$

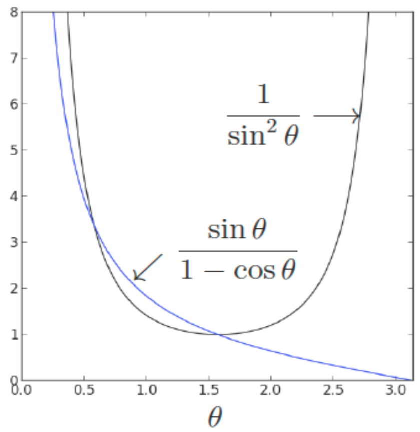

Figure 3. Consider the unit sphere $\mathbb{S}^2$ and a mass m_1 fixed at the north pole. The force on a second mass m_2 at a co-latitude θ is proportional to $1/\sin^2\theta$ if derived from the potential $\mathcal{U}(\theta) = -\tilde{\gamma}\, m_1 m_2 \cot\theta$, whereas it is proportional to $\sin\theta/(1-\cos\theta)$ if derived from the potential $\mathcal{U}(\theta) = \frac{\gamma}{4\pi}\, m_1 m_2 \ln(1-\cos\theta)$.

3) For a truly closed surface (i.e. compact and without boundaries or punctures), such as any surface conformal to the sphere [5,13,19,26], the source term $\gamma\rho$ of the Poisson equation $\Delta_g \Phi = \gamma\rho$ must satisfy an extra condition, called the *Gauss condition*:

$$\gamma \iint_M \rho d\boldsymbol{r} = 0. \tag{4.6}$$

This implies that the equation for the Green function must be generalized to

$$\Delta_g G(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') + C, \tag{4.7}$$

where C is a compensating factor chosen so that the surface integral of $\delta(\boldsymbol{r} - \boldsymbol{r}') + C$ is identically zero. C plays the role of a gauge.[1] The simplest choice is to take $C = -1/A$, a constant [5,13,26]. Then, C represents a uniform *background* distribution of matter which links the local dynamics to the global geometry of the surface. The fact that this distribution is negative could be interpreted as an anti-matter distribution (see [6,13,26] for further remarks). Notably, with such a choice for C, the Green function of a sphere of radius R is

$$G(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{4\pi} \ln[2R^2(1-\cos\Theta)]$$

where Θ is the angular separation between $\boldsymbol{r}$ and $\boldsymbol{r}'$. Note: $2R^2(1-\cos\Theta)$ is the chord distance between these two points. In G above, the radius of the sphere scales out — it contributes an

[1]It is important to stress that we are making a choice of *gauge* by choosing $C = 1/A$. Other choices are possible. For example : $a)$ for each mass m_j, $j = 1, \ldots, N$, choose an antipodal negative mass $m_{j+N} = -m_j$, where N is the total number of bodies in the system (cf. [25]); $b)$ a non constant compensating mass field; $c)$ a combination of $a)$ & $b)$.
To each one of those choices corresponds a different Green function. It would be very interesting to see, with an experiment, what would be the most natural choice. Our choice is, partially, motivated by the fact that with a compensating mass we would have to choose not only the the location but also the momentum of the corresponding negative mass.

unimportant constant. Dynamically, all spheres are equivalent to the unit sphere. The same is not true for the punctured sphere, as pointed out in remark 2) above.

4) It follows from remark 3) that on a surface M conformal to the sphere, we cannot just consider a single point-like mass, since in this case the integral (4.6) would not be zero. There are many options for compensating the mass. Among the simplest is to add an equal but opposite (negative) mass, as suggested by Shchepetilov [25]. But perhaps the simplest choice of all is to add a *uniform* background mass, of negative sign, because *this choice alone does not increase the dynamical degrees of freedom of the system* (cf. [13]). Thus, if we consider a point-like distribution of matter, comprised of N point-like masses, we must also consider a compensating term C_N:

$$\rho(\boldsymbol{r},t)=\sum_{j=1}^{N} m_j\delta(\boldsymbol{r}-\boldsymbol{r}_j(t))+C_N. \tag{4.8}$$

Here, we choose $C_N=-(\sum_{j=1}^{N} m_j)/A$. Then, inserting Eq. (4.8) into Eq. (4.4), we obtain the satellite potential Φ from

$$\Phi(\boldsymbol{r},t)=\gamma\sum_{j=1}^{N} m_j\, G(\boldsymbol{r},\boldsymbol{r}_j(t))+\text{constant}.$$

where G is the Green function found by solving Eq. (4.7) with $C=-1/A$.

5. The potential of N bodies on a surface (M,g)

The previous section explained how to deduce the satellite gravitational potential, assuming that the trajectories of the masses (the sources of the gravitational field) are known. Here we describe how to deduce the potential of the masses themselves. This relies on the following working hypothesis: *Each mass behaves as a satellite in the gravitational field generated by the other masses.*

A similar hypothesis is necessary to formulate the dynamics of point vortices in an incompressible fluid. Following [13], we define the potential of the kth mass as

$$\Phi(\boldsymbol{r}_k,t)=\lim_{\boldsymbol{r}\to\boldsymbol{r}_k}\left(\Phi(\boldsymbol{r},t)-\gamma\frac{m_k}{2\pi}\ln d(\boldsymbol{r},\boldsymbol{r}_k)\right)$$

where $d(\boldsymbol{r},\boldsymbol{r}_k)$ is the geodesic distance between $\boldsymbol{r}$ and $\boldsymbol{r}_k$. We thus obtain

$$\Phi(\boldsymbol{r}_k,t)=\gamma\Big(\sum_{j=1,\,j\neq k}^{N} m_jG(\boldsymbol{r}_k,\boldsymbol{r}_j)\;+\;m_k\mathcal{R}(\boldsymbol{r}_k)\Big), \tag{5.1}$$

where $\mathcal{R}(\boldsymbol{r}_k)$ is the Robin function [5,13,17],

$$\mathcal{R}(\boldsymbol{r}_k)=\lim_{\boldsymbol{r}\to\boldsymbol{r}_k}\left(G(\boldsymbol{r},\boldsymbol{r}_k)-\frac{1}{2\pi}\ln d(\boldsymbol{r},\boldsymbol{r}_k)\right).$$

It can then be inferred that the gravitational potential $\mathcal{U}$ takes the form

$$\mathcal{U}\;=\;\frac{1}{2}\iint_M\Phi(\boldsymbol{r},t)\rho(\boldsymbol{r},t)\,\mathrm{d}\Omega_{\boldsymbol{r}}.$$

For a system of N point masses, m_1, ..., m_N, located at positions $\boldsymbol{r}_1$, ..., $\boldsymbol{r}_N$, this reduces to

$$\begin{aligned}\mathcal{U}\;&=\;\frac{1}{2}\iint_M\Phi(\boldsymbol{r},t)\sum_{j=1}^{N} m_j\left(\delta(\boldsymbol{r}-\boldsymbol{r}_j)-\frac{1}{A}\right)\mathrm{d}\Omega_{\boldsymbol{r}},\\ &=\;\frac{1}{2}\sum_{j=1}^{N} m_j\Phi(\boldsymbol{r}_j,t)-\frac{m_{tot}}{2\,A}\iint_M\Phi(\boldsymbol{r},t)\mathrm{d}\Omega_{\boldsymbol{r}},\end{aligned} \tag{5.2}$$

where $m_{tot} = \sum_j m_j$ is used henceforth to denote the sum of the masses. By direct analogy with the (excess) energy of a system of point vortices [13], we can now give the explicit form of the gravitational potential energy.

Proposition 5.1. *The gravitational potential energy of a system of N point masses is*

$$\mathcal{U} = \gamma \sum_{k=1}^{N} \sum_{j<k}^{N} m_j m_k G(\boldsymbol{r}_j, \boldsymbol{r}_k) + \frac{\gamma}{2} \sum_{j=1}^{N} m_j^2 \mathcal{R}(\boldsymbol{r}_j) - \frac{\gamma\, m_{tot}}{2A} \sum_{j=1}^{N} m_j \iint_M G(\boldsymbol{r}, \boldsymbol{r}_j) d\Omega_{\boldsymbol{r}}. \tag{5.3}$$

The proof follows immediately from Eq. (5.1) and Eq. (5.2).

Remarks:

(a) The Robin function represents the self-interaction of a single mass with the global geometry of the surface. In particular, with the gauge choice made for the compensating term — i.e. the uniform background mass distribution — the Robin function is directly related to variations of the Gaussian curvature K over the surface M [26]:

$$\mathcal{R}(\boldsymbol{r}) = \frac{1}{2\pi} \iint_M G(\boldsymbol{r}, \boldsymbol{r}') K(\boldsymbol{r}') d\boldsymbol{r}' + c_1, \tag{5.4}$$

where c_1 is a constant. For surfaces with constant Gaussian curvature such as the sphere, it follows that the Robin function is itself a constant [13], and therefore plays no role in the mass dynamics.

(b) The expression (5.3) — as for the vortex Hamiltonian H (2.14) in [13] — holds for any closed, differentiable, genus zero surface (i.e. any surface topologically equivalent to a sphere).

(c) The last term in Eq. (5.3) does not contribute to the dynamics as the integral of Φ over the whole surface is a constant. We can then simplify the gravitational potential energy to

$$\mathcal{U} = \gamma \sum_{k=1}^{N} \sum_{j<k}^{N} m_j m_k G(\boldsymbol{r}_j, \boldsymbol{r}_k) + \frac{\gamma}{2} \sum_{j=1}^{N} m_j^2 \mathcal{R}(\boldsymbol{r}_j), \tag{5.5}$$

where the Green function part describes the gravitational interaction between pairs of masses, while the Robin function part can be viewed as the gravitational potential describing the interaction of a single mass with its uniform compensating mass distribution over the surface. As shown in the example of an ellipsoid — see §8 — it is through $\mathcal{R}$ that a single mass can still move on M. Explicit forms of $\mathcal{R}$ are given in [13].

6. The kinetic energy of a system of N point masses on (M, g)

It follows from the fact that we have a mechanical system whose phase space, mathematically speaking, is a cotangent bundle over the product manifold M^N, i.e. T^*M^N [2], that we can deduce the mass dynamics as follows.

Proposition 6.1. *Given a manifold M with metric g, the kinetic energy $\mathcal{K}$ of a mechanical system (i.e. a system which obeys Newton's second law) consisting of N point masses m_1, ..., m_N is*

$$\mathcal{K} = \sum_{j=1}^{N} \frac{1}{2m_j} \|\boldsymbol{p}_j\|_{g^{-1}}^2, \tag{6.1}$$

where $\|\boldsymbol{p}_j\|_{g^{-1}}^2 = \boldsymbol{p}_j^T (g^{-1})^T \boldsymbol{p}_j$, g^{-1} is the inverse of the 2×2 metric tensor g, and $\boldsymbol{p}_j$ is the momentum of the jth mass.

Proof. There is a simple way to prove the proposition above using the Lagrangian formulation and Legendre transformations [16] . The Lagrangian $\mathcal{L}$ of a system of particles of masses

$m_1, ..., m_N$ at positions, respectively, $\boldsymbol{r}_1, ..., \boldsymbol{r}_N$, moving at velocities $\boldsymbol{v}_1, ..., \boldsymbol{v}_N$ is

$$\begin{aligned} \mathcal{L} = \mathcal{K}(\boldsymbol{v}_1, ..., \boldsymbol{v}_N) - \mathcal{U}(\boldsymbol{r}_1, ..., \boldsymbol{v}_N) &= \frac{1}{2}\sum_{j=1}^{N} m_j \|\boldsymbol{v}_j\|_g^2 - \mathcal{U}(\boldsymbol{r}_1, ..., \boldsymbol{r}_N), \\ &= \frac{1}{2}\sum_{j=1}^{N} m_j {\boldsymbol{v}_j}^T g\, \boldsymbol{v}_j - \mathcal{U}(\boldsymbol{r}_1, ..., \boldsymbol{r}_N). \end{aligned}$$

The momentum $\boldsymbol{p}_j$ is defined by

$$\boldsymbol{p}_j = \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_j} = m_j {\boldsymbol{v}_j}^T g \quad \Longrightarrow \quad {\boldsymbol{v}_j}^T = \frac{1}{m_j} \boldsymbol{p}_j\, g^{-1} \quad \Longrightarrow \quad \boldsymbol{v}_j = (g^{-1})^T \boldsymbol{p}_j^T.$$

It follows that the kinetic energy can be expressed as

$$\mathcal{K} = \frac{1}{2}\sum_{j=1}^{N} m_j {\boldsymbol{v}_j}^T g\, \boldsymbol{v}_j = \sum j = 1^N \frac{1}{2m_j} \boldsymbol{p}_j\, (g^{-1})^T\, {\boldsymbol{p}_j}^T = \sum_{j=1}^{N} \frac{1}{2m_j} \boldsymbol{p}\, g^{-1}\, \boldsymbol{p}^T = \sum_{j=1}^{N} \frac{1}{2m_j} \|\boldsymbol{p}_j\|_{g^{-1}}^2,$$

where $g^T = g$ since g is a symmetric tensor. □

Example. For the sphere of radius R, where $\boldsymbol{r}_j = (\varphi_j, \theta_j)$, (with φ_j the longitude and θ_j the co-latitude), the momentum is $\boldsymbol{p}_j = (p_{j\,\varphi}, p_{j\,\theta})$. The metric is $\mathrm{d}s^2 = R^2[\mathrm{d}\theta^2 + \sin^2\theta\, \mathrm{d}\varphi^2]$, corresponding to which the metric tensor g and its inverse are given by

$$g = \begin{pmatrix} R^2 \sin^2\theta & 0 \\ 0 & R^2 \end{pmatrix} \qquad \text{and} \qquad g^{-1} = \begin{pmatrix} 1/(R^2 \sin^2\theta) & 0 \\ 0 & 1/R^2 \end{pmatrix}.$$

It follows from the proposition above that the kinetic energy is

$$\mathcal{K} \;=\; \sum_{j=1}^{N} \frac{1}{2m_j} (p_{j\,\varphi}, p_{j\,\theta}) \begin{pmatrix} 1/\sin^2\theta_j & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_{j\,\varphi} \\ p_{j\,\theta} \end{pmatrix} = \sum_{j=1}^{N} \frac{1}{2m_j\, R^2} \left(\frac{p_{j\,\varphi}^2}{\sin^2\theta_j} + p_{j\,\theta}^2 \right).$$

It follows from the proposition above that the configuration space of a system of N point masses is the product space M^N whose metric is $\mathrm{d}s_M^2 = \sum_{j=1}^N m_j ds_j^2$. The corresponding metric tensor is given by

$$\mathcal{G} = \begin{pmatrix} m_1 g & \dots & O \\ O & \ddots & O \\ O & \dots & m_N g \end{pmatrix}$$

where O is a 2×2 matrix of zeros. Moreover, the motion of the point masses is described by the Hamiltonian system

$$\dot{\boldsymbol{r}}_j = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_j}, \quad \dot{\boldsymbol{p}}_j = \frac{\partial \mathcal{H}}{\partial \boldsymbol{r}_j}, \qquad j = 1, ..., N, \tag{6.2}$$

where

$$\mathcal{H}(\boldsymbol{Q}, \boldsymbol{P}) = \frac{\|\boldsymbol{P}\|_{\mathcal{G}^{-1}}^2}{2} + \mathcal{U}(\boldsymbol{Q}), \tag{6.3}$$

with $\boldsymbol{P} = (\boldsymbol{p}_1, ..., \boldsymbol{p}_N)$, $\boldsymbol{Q} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$, $\|\boldsymbol{P}\|_{\mathcal{G}^{-1}}^2 = \boldsymbol{P}^T \mathcal{G}^{-1} \boldsymbol{P}$ and $\mathcal{U}(\boldsymbol{Q}) = \mathcal{U}(\boldsymbol{r}_1, \dots, \boldsymbol{r}_n)$ is the potential in Eq. (5.5).

Remark: There is a difference between a satellite and a passive tracer advected by vortices in an incompressible fluid. The passive tracer genuinely has zero vorticity, whereas the satellite has a small, but non-zero, mass. So taking this into account, the satellite potential has a small (perturbative) self-interaction through the Robin function:

$$\Phi_0(\boldsymbol{r},t) = \gamma \sum_{j=1}^{N} m_j G(\boldsymbol{r}_0, \boldsymbol{r}_j(t)) + \gamma\, m_0 \mathcal{R}(\boldsymbol{r}_0).$$

with $m_0 \ll m_j$, $j = 1, \ldots, N$. The satellite kinetic energy is $\mathcal{K}_0 = \|\boldsymbol{p}\|^2_{g^{-1}}/(2\,m_0)$ (by Proposition 6.1), and thus the satellite Hamiltonian is $\mathcal{H}_0(\boldsymbol{r},\boldsymbol{p},t) = \mathcal{K}_0 + m_0 \Phi_0(\boldsymbol{r},t)$. Mathematically speaking, the configuration space is the surface M with metric g, and the phase space is the cotangent bundle T^*M.

7. Symmetries and reduction

When studying the dynamics of a N masses, the process of reduction makes use of the following fundamental symmetries:

a) **The symmetries of the Hamiltonian**. The groups of transformations that leave the Hamiltonian invariant. Such groups depend on the surface geometry. In the case of the plane they are the group of translations and the group of rotations, while in the case of the sphere we only have the group of rotations.
b) **The symmetries of the equations**. The group of time transformations that leaves the equations of motion invariant. This group depends both on the degree of *separability* of the Hamiltonian and on the Robin function. We define a Hamiltonian as separable when it can be viewed as the sum of two distinct functions.

(a) Symmetries in common to all surfaces of revolution conformal to $\mathbb{S}^2$

Consider the N-body problem on a surface of revolution for which the masses are parametrised by $\boldsymbol{r}_j = (\varphi_j, \theta_j)$, $j = 1, \ldots, N$,. Let $\mathrm{d}s^2 = f(\theta)\mathrm{d}\varphi^2 + \sigma(\theta)\mathrm{d}\theta^2$ be the surface metric. Then it follows from Proposition 6.1 that the corresponding kinetic energy is

$$\mathcal{K} = \sum_{j=1}^{N} \frac{1}{2m_j}\left(\frac{p_{j\,\varphi}^2}{f(\theta_j)} + \frac{p_{j\,\theta}^2}{\sigma(\theta_j)}\right) \tag{7.1}$$

Thus the kinetic energy does not depend upon the longitudes φ_j, $j = 1, \ldots, N$, as expected due to the axial symmetry of the surface geometry. Furthermore, the potential energy, see Eq. (5.5) and [13], can be decomposed as

$$\mathcal{U} = \gamma \sum_{j=1}^{N}\sum_{k>j}^{N} m_j m_k G(|\varphi_j - \varphi_k|, \theta_j, \theta) + \gamma \sum_{j=1}^{N} m_j^2 \mathcal{R}(\theta_j). \tag{7.2}$$

Therefore, by Noether's theorem [2], we have conservation of angular momentum, $P_\varphi = \sum_{j=1}^{N} p_{j\,\varphi}$. In addition, for the pair of canonical variables $(\varphi_j, p_{j\,\varphi})$, $j = 1, \ldots, N$, the Hamiltonian equations of motion simplify to

$$\dot{\varphi}_j = \frac{\partial \mathcal{K}}{\partial p_{j\,\varphi}}, \qquad \dot{p}_{j\,\varphi} = -\frac{\partial \mathcal{U}}{\partial \varphi_j}, \qquad j = 1, \ldots, N,$$

which gives

$$\ddot{\varphi}_j = -\frac{\partial \mathcal{U}}{\partial \varphi_j}, \qquad j = 1, \ldots, N.$$

Using Eq. (7.2), it can be verified that the equations above are invariant with respect to the time-varying coordinate transformation

$$\tilde{\varphi}_j = \varphi_j + \nu\, t.$$

All of the above can be then summarized in the following proposition

Proposition 7.1. *For a system of N-bodies on a surface of revolution, the total φ component of the momentum, $P_\varphi = \sum_{j=1}^N p_{j\,\varphi}$, is an integral of motion.*
Furthermore the equations of motion are invariant under the time-dependent transformation

$$\tilde{\varphi}_j = \varphi_j + \nu\, t \qquad \forall \nu \in \mathbb{R}. \tag{7.3}$$

Remarks

- Notice that the time-varying symmetry group (7.3) for surfaces of revolution directly corresponds to the Galilean group for $\mathbb{R}^2$, see Appendix B and [2].
- The group of transformations (7.3) leaves the equations of motion invariant but not the Hamiltonian. From Eq. (7.1) we have

  $$\dot{\varphi}_j = \frac{\partial \mathcal{K}}{\partial p_{j\,\varphi}} = \frac{p_{j\,\varphi}}{m_j f(\theta_j)},$$

  It follows that, under the transformation (7.3), the φ component of the momentum changes to

  $$\tilde{p}_{j\,\tilde{\varphi}} = m_j f(\theta_j)(\dot{\varphi}_j + \nu),$$

  which in turn changes the kinetic energy (7.1).
  Remark: Observe that in Euclidean spaces such as $\mathbb{R}^2$ and $\mathbb{R}^3$, the reduction procedure above can be viewed as introducing a ficticious center of mass on the surface M, though the real configuration space is M^N. On a general surface, we can no longer use the notion of center of mass, at least not as in Euclidean spaces (since it won't generally lie on M). Nevertheless the above coordinate transformations are legitimate on the configuration space M^N.

(b) The unit sphere $\mathbb{S}^2$

Consider a system of N masses, $m_1, m_2, ..., m_N$, at the positions $\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N$ on the unit sphere $\mathbb{S}^2$. In spherical coordinates $\boldsymbol{r}_j = (\varphi_j, \theta_j)$, $j = 1, ..., N$, the metric of the configuration space is

$$\mathrm{d}s^2 = \sum_{j=1}^N m_j \,(\sin^2\theta_j \mathrm{d}\varphi_j^2 + \mathrm{d}\theta_j^2).$$

The Green Function for the unit sphere problem [4,13,14,19] is

$$G(\boldsymbol{r}_j, \boldsymbol{r}_k) = G(|\varphi_j - \varphi_k|, \theta_j, \theta_k) = \frac{1}{4\pi}\log(1 - d_{jk})$$

where

$$d_{jk} = \cos\theta_j \cos\theta_k + \sin\theta_j \sin\theta_k \cos(\varphi_j - \varphi_k).$$

It follows from Eqs. (6.1) and (5.5) that the Hamiltonian is only partially separable and

$$H = \sum_{j=1}^N \frac{1}{m_j}\left(\frac{p_{j\,\varphi}^2}{\sin^2\theta_j} + p_{j\,\theta}^2\right) + \gamma \sum_{j=1}^N \sum_{k>j}^N m_j m_k G(|\varphi_j - \varphi_k|, \theta_j, \theta_k), \tag{7.4}$$

where γ is the gravitational constant of the unit sphere and, as for the plane, the Robin function is a constant [13]. The corresponding equations of motion are

$$\dot{\varphi}_j = \frac{\partial H}{\partial p_{j\,\varphi}}, \qquad \dot{\theta}_j = \frac{\partial H}{\partial p_{j\,\theta}}, \qquad \dot{p}_{j\,\varphi} = -\frac{\partial H}{\partial \varphi_j}, \qquad \dot{p}_{j\,\theta} = -\frac{\partial H}{\partial \theta_j}. \tag{7.5}$$

Note, the Hamiltonian is invariant under

- rotations with respect the three axes which gives the conservation of the total angular momentum $\boldsymbol{L} = (L_x, L_y, L_z)$; those three integrals do not all commute.

- time translations, which is related to the conservation of the Hamiltonian H (or excess energy [13]).

In addition, due to the partial separability of the Hamiltonian, as stated in Proposition 7.1 we have the group of symmetries (7.3)

$$\tilde{\varphi}_j = \varphi_j + \nu\, t, \qquad\qquad j = 1, ..., N,$$

i.e. time-dependent transformations which leave equations (7.5) invariant.

The symmetry of the sphere also leads to a simple Cartesian-coordinate formulation of the equations of motion. In Cartesian coordinates centred at the origin of the sphere, each position vector is a unit vector: $|\boldsymbol{r}_j| = 1$, $j = 1,\, 2,\, ...,\, N$. Using $\boldsymbol{r}_j = (\sin\theta_j \cos\varphi_j, \sin\theta_j \sin\varphi_j, \cos\theta_j)$ in (7.5), we find after some manipulation,

$$\begin{aligned}\dot{\boldsymbol{r}}_k &= \boldsymbol{u}_k\\ \dot{\boldsymbol{u}}_k &= \left(\tilde{M} - \tilde{m}_k - |\boldsymbol{u}_k|^2\right)\boldsymbol{r}_k + \gamma \sum_{j\neq k} \tilde{m}_j \frac{\boldsymbol{r}_j - \boldsymbol{r}_k}{1 - \boldsymbol{r}_j \cdot \boldsymbol{r}_k}\end{aligned}$$

where $\tilde{m}_k = m_k/4\pi$ and $\tilde{M} = \sum_{k=1}^{n} \tilde{m}_k$. One can verify that the identity $d(\boldsymbol{r}_k \cdot \boldsymbol{u}_k)/dt = \dot{\boldsymbol{r}}_k \cdot \boldsymbol{u}_k + \boldsymbol{r}_k \cdot \dot{\boldsymbol{u}}_k = 0$ is satisfied, as required.

The conserved total (kinetic plus potential) energy E is obtained from

$$E/4\pi = \frac{1}{2}\sum_{k=1}^{n} \tilde{m}_k |\boldsymbol{u}_k|^2 + \gamma \sum_{k=2}^{n}\sum_{j=1}^{k-1} \tilde{m}_j \tilde{m}_k \ln(1 - \boldsymbol{r}_j \cdot \boldsymbol{r}_k)$$

The two body problem. In the case of two masses, the corresponding system of equations (7.5) has eight degrees of freedom (four Hamiltonian degrees of freedom). Using conservation of L_x and L_y, together with the freedom in choosing the orientation of the z-axis (which amounts to choosing $L_x = L_y = 0$), the further conservation of L_z and the time transformation (7.3) reduce the original system to a system having four degrees of freedom. Furthermore using the conservation of Hamiltonian H itself, we can further reduce this to three. To assure integrability [2], we need one additional integral of motion.

Following the analysis of Manuele Santoprete [23], Rodrigo Schaefer proved the following theorem [24]:

Theorem 7.1. *Consider the Kepler problem on the unit sphere $\mathbb{S}^2$. One body of mass m_2 is held fixed at the north pole and the dynamics of the second body, of mass m_1 and position $\boldsymbol{r} = (\varphi, \theta)$, is described by the Hamiltonian equations*

$$\dot{\varphi} = \frac{\partial H}{\partial p_\varphi}, \quad \dot{\theta} = \frac{\partial H}{\partial p_\theta}, \quad \dot{p_\varphi} = -\frac{\partial H}{\partial \varphi} = 0, \quad \dot{p_\theta} = -\frac{\partial H}{\partial \theta},$$

where θ represents also the geodesic distance between the two bodies and

$$H(\theta, p_\varphi, p_\theta) = \frac{p_\varphi^2}{2m_1 \sin^2\theta} + \frac{p_\theta^2}{2m_1} + \gamma m_1 m_2 \log[1 - \cos(\theta)]$$

The only integrals which are quadratic in the momentum components, i.e. integrals of the type

$$I = a(\varphi, \theta)p_\theta^2 + 2b(\varphi, \theta)p_\varphi p_\theta + c(\varphi, \theta)p_\varphi^2 + g(\varphi, \theta),$$

are

$$I = 2\, m_1\, C_1\, H + C_2 p_\varphi^2\, + C_2, \qquad \forall C_1,\, C_2 \in \mathbb{R}.$$

Remarks:

- The extra integrals encountered in the theorem above are linear combinations of known integrals (plus a constant), and therefore they do not provide us with a truly new extra integral, as in the corresponding planar problem, see Appendix B, where the corresponding theorem (Theorem B.1) provides us with the components of the Laplace-Runge-Lenz vector.
- In the theorem above the dynamics is viewed from within the sphere's intrinsic geometry, while in Santoprete [23] the dynamics is viewed as a sphere embedded in $\mathbb{R}^3$. Consequently the potential Φ is different from the intrinsic one, as discussed above in §4.

8. Dynamics of one mass

For illustration, let us consider a single mass on a ellipsoid of revolution M, with $R=1$. As discussed in §3, the metric of M is $\mathrm{d}s^2 = (\cos^2\theta + b^2\sin^2\theta)\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2$, corresponding to which the metric tensor g and its inverse are given by

$$g = \begin{pmatrix} \sin^2\theta & 0 \\ 0 & (\cos^2\theta + b^2\sin^2\theta) \end{pmatrix} \quad \text{and} \quad g^{-1} = \begin{pmatrix} 1/\sin^2\theta & 0 \\ 0 & 1/(\cos^2\theta + b^2\sin^2\theta) \end{pmatrix}.$$

It follows from Proposition 6.1 that the configuration space of a system of one point mass M has the metric $\mathrm{d}s_M^2 = m\mathrm{d}s^2$. The corresponding metric tensor is given by $\mathcal{G} = m\,g$. As in Eq. (6.2), the motion of one point mass is described by the Hamiltonian system

$$\dot{\boldsymbol{r}} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}}, \quad \dot{\boldsymbol{p}} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{q}}, \tag{8.1}$$

where $\boldsymbol{r} = (\varphi, \theta)$ and $\boldsymbol{p} = (p_\varphi, p_\theta)$, together with

$$\mathcal{H}(\boldsymbol{q}, \boldsymbol{p}) = \frac{\|\boldsymbol{p}\|^2_{\mathcal{G}^{-1}}}{2} + \mathcal{U}(\boldsymbol{r}),$$

where

$$\|\boldsymbol{p}\|^2_{\mathcal{G}^{-1}} = \boldsymbol{p}^T\mathcal{G}^{-1}\boldsymbol{p} = \frac{p_\varphi^2}{2m\,\sin^2\theta} + \frac{p_\theta^2}{2m\,(\cos^2\theta + b^2\sin^2\theta)}$$

and

$$\mathcal{U}(\boldsymbol{r}) = \frac{\gamma\,m^2}{2}\mathcal{R}(\boldsymbol{r})$$

from Eq. (5.5). Here γ is the gravitational constant for the ellipsoid and $\mathcal{R}(\boldsymbol{r})$ is the Robin function [17], a pure function of θ for a surface of revolution [13].

Then from Eq. (8.1), a single mass on an ellipsoid evolves according to

$$\dot{\varphi} = \frac{\partial \mathcal{H}}{\partial p_\varphi} = \frac{p_\varphi}{m\sin^2\theta}, \qquad \dot{\theta} = \frac{\partial \mathcal{H}}{\partial p_\theta} = \frac{p_\theta}{m(\cos^2\theta + b^2\sin^2\theta)},$$

$$\dot{p}_\varphi = -\frac{\partial \mathcal{H}}{\partial \varphi} = 0, \qquad \dot{p}_\theta = -\frac{\partial \mathcal{H}}{\partial \theta} = \frac{p_\theta^2\,(b^2-1)\sin 2\theta}{2m(\cos^2\theta + b^2\sin^2\theta)^2} + \frac{p_\varphi^2\,\cos\theta}{m\sin^3\theta} - \frac{\gamma\,m^2}{2}\frac{d\mathcal{R}}{d\theta},$$

An explicit form for $d\mathcal{R}/d\theta$ may be found by combining Eqs. (4.18) and (5.10) in [13], giving

$$\frac{d\mathcal{R}}{d\theta} == -\frac{1}{2\pi\sin\theta}\left(\cos\theta - q(\theta)\frac{\mu(\theta)}{\mu(0)}\right)$$

where $q(\theta) = \sqrt{\cos^2\theta + b^2\sin^2\theta}$ and

$$\mu(\theta) = \frac{q\cos\theta}{2} + \begin{cases} \frac{b^2}{2\sqrt{1-b^2}}\ln\left(\frac{q+\sqrt{1-b^2}\cos\theta}{b}\right) & : b<1 \\ \frac{b^2}{2\sqrt{b^2-1}}\sin^{-1}\left(\frac{\sqrt{b^2-1}\cos\theta}{b}\right) & : b>1 \end{cases} \tag{8.2}$$

Note $4\pi\mu(0)$ gives the total surface area A of the ellipsoid.

As shown in Figure 2 (for $b = 0.5$), a mass initially at rest starts moving along a meridian ($\varphi =$ constant). This is caused by the interaction with the uniform negative mass spread over its surface, as expressed through the Robin function.

In particular, if the mass is initially at rest $p_\varphi(0) = p_\theta(0) = 0$, and located at $\varphi(0) = \varphi_0$, $\theta(0) \neq \{0,\ \frac{\pi}{2},\ \pi\}$, then the motion is purely along a meridian. This in fact occurs on any surface of revolution — see [6] and [13] for more details.

9. The two body problem and Kepler's laws

From the intrinsic geometry point of view, we observe that in the plane, $\mathbb{R}^2$, the two body potential becomes

$$\mathcal{U}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \gamma\, m_1 m_2 G_{\mathbb{R}^2}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{\gamma\, m_1\, m_2}{2\pi} \ln ||\boldsymbol{r}_1 - \boldsymbol{r}_2||.$$

where $\boldsymbol{r}_1 = (x_1, y_1)$ and $\boldsymbol{r}_2 = (x_2, y_2)$ [13,17]. As proved in Appendix B the two body problem is integrable. Nevertheless, as discussed in [2,24], Kepler's laws are no longer valid in such a geometry, as the two body problem admits a unique elliptic orbit, the circular one. On the other hand we have that all orbits are limited. Therefore Kepler's laws are no longer valid on the plane when viewed from the intrinsic geometry point of view — as opposed to planar motion embedded in $\mathbb{R}^3$.

Conclusions

We have shown how to formulate the dynamics of point masses on closed surfaces. A key aspect of the analysis is to properly account for the mathematical requirement that the mass integrated over the surface must vanish. This leads to significant differences from previous formulations, developed for punctured surfaces (not truly closed). For example, on a sphere, the radius of the sphere scales out of the gravitational potential, but on a punctured sphere it does not. This has a profound influence on the resulting equations of motion.

An interesting feature of our formulation of point mass dynamics on closed surfaces is that Newton's famous Law of Inertia does not hold generally (on any surface with variable Gaussian curvature). This means that a particle at rest can begin to move. This is caused by the interaction with the geometry, specifically with the uniform (negative) mass spread uniformly over the surface. Moreover, Kepler's laws, originally formulated for three-dimensional space, generally do not hold on closed surfaces. They do not hold even for two-dimensional motion on the plane, even though they do hold for planar motion in three-dimensional space.

Finally, and perhaps unexpectedly, variations of Gaussian curvature on closed surfaces generate dynamics. This may be viewed as the classical analogue of the Equivalence Principle of General Relativity, where the curvature of space-time is equivalent to a force field.

Acknowledgements. The authors are grateful to Florin Diacu for his Lectures on the curved N-body problem at UFRJ and IMPA in October/November of 2012 which motivated this work. The authors would like to thank Umberto Hryniewicz, Richard Montgomery, Carles Simó and Teresa Stuchi for helpful discussions.

Funding statement. Rodrigo Schaefer was supported by FAPERJ (Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro) through the fellowship FAPERJ Nota 10. SB acknowledges fundings from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPERJ. DGD gratefully acknowledges support for this research from CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) and FINEP (Inovação e Pesquisa) in Brazil, and from the UK Engineering and Physical Sciences Research Council (grant no. EP/H001794/1).

A. Hodge decomposition

We begin by discussing the Hodge decomposition theorem for vector fields in $\mathbb{R}^2$ and $\mathbb{R}^3$. We closely follow Baird [3] and Chorin and Marsden [9]. Any well-behaved (at least twice differentiable) vector field $\boldsymbol{v}$ on $\mathbb{R}^2$ or $\mathbb{R}^3$ can be decomposed in three components: *transverse* $\boldsymbol{v}_1$ (rotational and non-divergent), *radial* $\boldsymbol{v}_2$ (irrotational and divergent), and *Laplacian* $\boldsymbol{v}_3$ (irrotational and non-divergent), i.e.

$$\boldsymbol{v} \quad = \quad \boldsymbol{v}_1 \quad + \quad \boldsymbol{v}_2 \quad + \quad \boldsymbol{v}_3,$$

with

$$div(\boldsymbol{v}_1) = 0, \qquad div(\boldsymbol{v}_2) = \Lambda \qquad div(\boldsymbol{v}_3) = 0, \tag{A 1}$$

$$curl(\boldsymbol{v}_1) = \Sigma, \qquad curl(\boldsymbol{v}_2) = 0, \qquad curl(\boldsymbol{v}_3) = 0. \tag{A 2}$$

In $\mathbb{R}^2$, using Cartesian coordinates, equations $div(\boldsymbol{v}_1) = 0$ and $curl(\boldsymbol{v}_2) = 0$ permit one to introduce, respectively, the functions Ψ and Φ such that the components $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ of the vector field can be written as

$$\boldsymbol{v}_1 = J\nabla\Psi, \qquad \boldsymbol{v}_2 = -\nabla\Phi, \tag{A 3}$$

where J is the usual co-symplectic matrix $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})^T$ is the usual gradient.

To determine Ψ and Φ we then substitute Eqs. (A 3) into the equations $curl(\boldsymbol{v}_1) = \Sigma$ and $div(\boldsymbol{v}_2) = \Lambda$ to obtain

$$\Delta\Psi = \Sigma, \qquad \Delta\Phi = -\Lambda.$$

Remarks

1) Observe that vector fields belonging to the third class, i.e. verifying

$$div(\boldsymbol{v}_3) = 0 \quad \text{and} \quad curl(\boldsymbol{v}_3) = 0,$$

can be expressed as $\boldsymbol{v}_3 = -\nabla\Phi_h$ where Φ_h is a harmonic function, i.e. satisfying $\Delta\Phi_h = 0$.

2) The solution of the Poisson equation $\Delta\Phi = -\Lambda$ is not unique since we can always add a solution of the corresponding harmonic equation $\Delta\Phi_h = 0$. To restrict the class of harmonic functions to constant functions, extra conditions are necessary. For vector fields in $\mathbb{R}^2$ the *extra condition* is to require that far from any sources, the velocity field tends to zero ($\boldsymbol{v}(\boldsymbol{r}) \longrightarrow 0$ as $|\boldsymbol{r}| \to \infty$). In general, each surface has its own set of extra conditions but these are not known in general.

3) There is more general version of the Hodge decomposition theorem that can be used for any simply connected domain D of $\mathbb{R}^2$, as well as for simply-connected surfaces such as the sphere, surfaces conformal to the sphere, and the hyperbolic plane. This generalised theorem is called the "one form decomposition theorem" (see [7,12,24,28]). Given a surface M with metric g, we can always associate a one form field to a given vector field by using the metric tensor. In fact if $\boldsymbol{a}$ is a vector field on M the corresponding one form field is $\sigma = g(\boldsymbol{a}, \cdot)$ — or equivalently $\sigma = \boldsymbol{a}^T g$ in local coordinates.

4) Let $\boldsymbol{v} = \boldsymbol{u}$ be the velocity field of a fluid particle — also called a *passive tracer*. A fluid is said to be *incompressible* if

$$div(\boldsymbol{u}) = 0.$$

Given the vorticity field defined as $\omega = curl(\boldsymbol{u})$, incompressible fluids belong to the first class of vector fields above. For the class of surfaces we are considering, using the coordinates of the area form (see [4,13]), the first equation allows us to re-write the velocity field as $\boldsymbol{u} = J\boldsymbol{\nabla}\Psi$, where Ψ is a suitably regular function, called the streamfunction, to be determined. Substituting the equation above into $curl(\boldsymbol{u}) = \omega$, we obtain $\Delta_g\Psi = \omega$, which is the equation that defines Ψ for a given metric g and a given vorticity field ω. For more details see [4,13].

5) Let $\boldsymbol{v} = \boldsymbol{a}$ be the acceleration field of a test particle of mass m_o. The corresponding force field $\boldsymbol{F} = m_o\boldsymbol{a}$ is a *central force* field if

$$curl(\boldsymbol{a}) = 0$$

and $div(\boldsymbol{a}) = \pm\gamma\rho$, where γ is a constant, and ρ is the density of the sources generating the central force. The plus or minus sign in $div(\mathbf{a}) = \pm\gamma\rho$ refers to the fact that we can have either a repulsive or attractive force. It follows that we can re-write $\boldsymbol{a} = -\nabla\Psi$ where Ψ is solution of $\Delta_g\Psi = \mp\gamma\rho$.

B. The planar problem: symmetries and extra integrals

Consider a system of N masses, m_1, m_2, ..., m_N with corresponding positions $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, ..., $\boldsymbol{r}_N$ on the plane $\mathbb{R}^2$. We consider Cartesian coordiantes, $\boldsymbol{r}_j = (x_j, y_j)$, $j = 1, ..., N$. The metric of the configuration space is then simply

$$\mathrm{d}s^2 = \sum_{j=1}^{N} m_j\,(\mathrm{d}x_j^2 + \mathrm{d}y_j^2).$$

The Green Function for the planar problem [4,13] is

$$G(\boldsymbol{r}_j, \boldsymbol{r}_k) = G(\|\boldsymbol{r}_j - \boldsymbol{r}_k\|) = \frac{1}{4\pi}\log\|\boldsymbol{r}_j - \boldsymbol{r}_k\|^2$$

It follows from Eqs. (6.1) and (5.5) that the Hamiltonian is separable and

$$H = \sum_{j=1}^{N}\frac{1}{m_j}(p_{j\,x}^2 + p_{j\,y}^2) + \gamma\sum_{j=1}^{N}\sum_{k>j}^{N} m_j m_k G(\|\boldsymbol{r}_j - \boldsymbol{r}_k\|),$$

since in this case the Robin function is a constant [13,17].
Note, the Hamiltonian is invariant under

- rotations with respect the axis perpendicular to the plane, which implies conservation of total angular momentum L;
- translations with respect to the x and y axis, which implies conservation of the total linear momentum $P = (P_x, P_y)$, where $P_x = \sum_{j=1}^{N} p_{j,\,x},\;\; P_y = \sum_{j=1}^{N} p_{j,\,y}$;
- time translations, which implies conservation of the Hamiltonian H (excess energy).

Due to the separability of the Hamiltonian — i.e. $H = \mathcal{K}(\boldsymbol{P}) + \mathcal{U}(\boldsymbol{Q})$ with $\boldsymbol{Q} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$ and $\boldsymbol{P} = (\boldsymbol{p}_1, ..., \boldsymbol{p}_N)$ — the Hamiltonian equations (6.2) simplify to

$$\dot{\boldsymbol{r}}_j = \frac{\partial\mathcal{K}}{\partial\boldsymbol{p}_j}, \quad \dot{\boldsymbol{p}}_j = -\frac{\partial\mathcal{U}}{\partial\boldsymbol{r}_j}, \tag{A 1}$$

and the first-order system of equations above is equivalent to the second-order system

$$\ddot{\boldsymbol{r}}_j = -\frac{\partial\mathcal{U}}{\partial\boldsymbol{r}_j}, \qquad j = 1, ..., N. \tag{A 2}$$

It follows that the equations of motion have an additional symmetry property: they are invariant with respect to the time-varying coordinate transformations — the Galilean group [2] —

$$\tilde{x}_j = x_j + v_x t, \qquad \tilde{y}_j = y_j + v_y t, \qquad j = 1, ..., N.$$

(a) The two body problem and the Laplace-Runge-Lenz vector

In the case of two masses, the corresponding system of equations (A 1) has eight degrees of freedom (four Hamiltonian degrees of freedom). Using conservation of linear momentum $\boldsymbol{P} = (P_x, P_y)$ and the Galilean group above, the original system reduces to a system of four degrees of freedom. Furthermore using the invariance under rotations and fixing the angular momentum L to a specific value, we further reduce this to two degrees of freedom. Finally, using the conservation of the Hamiltonian H, the system is reduced to a single degree of freedom. The two body problem in the plane is therefore integrable.

Remark. Following the analysis of Manuele Santoprete [23], Rodrigo Schaefer proved the following theorem [7,24]:

Theorem B.1. *Consider the Kepler problem on the plane* $\mathbb{R}^2$*. One body of mass* m_2 *is held fixed at the origin of a chosen reference system and the dynamics of the second body, has, respectively, mass* m_1 *and position* $\boldsymbol{r} = (r, \varphi)$*, where* r *and* φ *are the usual polar coordinates. The Kepler problem is described by the Hamiltonian equations*

$$\dot{\varphi} = \frac{\partial H}{\partial p_\varphi}, \quad \dot{r} = \frac{\partial H}{\partial p_r}, \quad \dot{p}_\varphi = -\frac{\partial H}{\partial \varphi} = 0, \quad \dot{p}_r = -\frac{\partial H}{\partial r},$$

where r *represents also the geodesic distance between the two bodies and*

$$H(r, p_\varphi, p_r) = \frac{p_\varphi^2}{2m_1 r^2} + \frac{p_r^2}{2m_1} + \gamma m_1 m_2 \log(r)$$

The only integrals which are quadratic in the momentum variables, i.e. integrals of the type

$$I = a(r, \phi)p_r^2 + 2b(r, \varphi)p_\varphi p_r + c(r, \varphi)p_\varphi^2 + \chi(r, \varphi),$$

are

$$I_1 = \sin\varphi\, p_r p_\varphi + \frac{\cos\varphi}{r} p_\varphi^2 - \gamma m_1 m_2 \cos\varphi, \tag{A 3}$$

$$I_1 = -\cos\varphi\, p_r p_\varphi + \frac{\sin\varphi}{r} p_\varphi^2 - \gamma m_1 m_2 \sin\varphi. \tag{A 4}$$

Remarks

- In the literature, the vector $I = (I_1, I_2)$ is also called the Laplace-Runge-Lenz vector [16].
- [7,24] proved that three of the four integrals H, p_φ, I_1 and I_2 are independent integrals and in involution. Therefore the components of the Laplace-Runge-Lenz vector provide us with an extra integral in the case of planar dynamics.

References

1. Acker F. 2013 *Depto Matemática Aplicada, Instituto de Matemática, Universidade Federal de Rio de Janeiro.* Felipe Acker is the one who suggested *Newton axioms* for Newton Laws.
2. Arnold VI. 1997 *Mathematical Methods of Classical Mechanics*. Springer.
3. See C.S. Baird website http://faculty.uml.edu/cbaird/95.657(2012)/Helmholtz_Decomposition.pdf
4. Boatto S. 2008. Curvature perturbation and stability of a ring of vortices. *Discrete and Continuous Dynamical Systems Series B* **10**, Nos. 2 & 3, 349–375.
5. Boatto S, Koiller J. 2015. *Vortices on Closed Surfaces*, published in *Geometry, Mechanics and Dynamics: the Legacy of Jerry Marsden.*, Fields Institute Communications series, **73**, 185–237. Springer.
6. Boatto S, Schaefer R, Dritschel D. 2016. N-vortex and N-body dynamics on closed surfaces 1. A classical equivalence principle. In preparation.
7. Boatto S, Dos Santos Loiola R, Schaefer R, Duarte G, Stuchi T, Dritschel D. Dynamic of masses on surfaces: a geometry point of view. In preparation.
8. Borisov AV, Mamaev IS, Kilin AA. 2004. Two-body problem on a sphere, reduction, stocasticity, periodic orbits. *Regular and Chaotic Dynamics* **9**, No. 3, 265–279.
9. Chorin, A. and Marsden, J. E. , 2000. *A Mathematical Introduction to Fluid Mechanics.* Springer-Verlag
10. Diacu F. 2012. Relative Equilibria of the Curved N-Body Problem. *Atlantis Studies in Dynamical Systems*, Vol. 1. Amsterdam: Atlantis Press.
11. Do Carmo MP. 1976 *Differential geometry of curves and surfaces*. Prentice-Hall.

12. Dos Santos Loiola R. 2016. *Dinâmica de vórtices e massas sobresuperfícies hiperbolicas*. Master thesis in Mathematics, Federal University of Rio de Janeiro, Brazil
13. Dritschel DG, Boatto S. 2015. The motion of point vortices on closed surfaces. *Proc. R. Soc. A* **471**, 20140890.
14. Dritschel, DG. 1988. Contour dynamics/surgery on the sphere. *J. Comput. Phys.* **78**, 477–483.
15. Evans LC. 2010 *Partial Differential Equations*. Graduate Studies in Mathematics, Vol. 19. Amer. Math. Soc.
16. Goldstein H, Poole CP, Safko JL. 2014 *Classical Mechanics*. Pearson New International Edition.
17. Flucher M, Gustafsson B. 1997. Vortex motion in two-dimensional hydrodynamics, energy renormalization and stability of vortex pairs. Preprint. See https://people.kth.se/ gbjorn/vortex.pdf.
18. Jackson JD. 1999. *Classical electrodynamics*. Wiley.
19. Kimura Y. 1999. Vortex motion on surfaces with constant curvature. *Proc. R. Soc. Lond. A* **455**, 245–259.
20. Koslov VV, Harin AO. 1992. Kepler's problem in constant curvature space. *Celestial Mechanics and Dynamical Astronomy* **54**, 393–399.
21. Newton I. 1686. *PhilosophiæNaturalis Principia Mathematica*. Springer. See http://cudl.lib.cam.ac.uk/view/PR-ADV-B-00039-00001/1.
22. Poincaré H. 1899. *Les méthodes nouvelles de la mécanique*. Gauthier-Villars et fils.
23. Santoprete M. 2009. Block regularization of the Kepler problem on surfaces of revolution with positive constant curvature. *Journal of Differential Equations* **247**, Issue 4, 1043–1063.
24. Schaefer R. 2014 *A dinâmica de N corpos sobre superfícies: um exemplo para 2 corpos*. Master thesis in Mathematics, Federal University of Rio de Janeiro, Brazil
25. Shchepetilov AV. 2006. *Calculus and Mechanics on two points homogeneous Riemannian Spaces*. Lecture Notes in Physics, **707**. Springer.
26. Steiner J. 2003. *Green's Functions, Spectral Invariants, and a Positive mass on Spheres*. PhD thesis in Mathematics, University of California, San Diego.
27. Steiner J. 2005. A geometrical mass and its extremal properties for metrics on S^2. *Duke Math. J.* **129**, 63–86.
28. Warner, F. W. , 1983. *Foundation of Differentiable Manifolds and Lie Groups*. Graduate Texts in Mathematics, Springer-Verlag.