© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Chapter One

Introduction to the N-Body Problem

The basic differential equations are defined that we will use throughout this book. These include the Newtonian n-body problem in section 1.1, and the planar three-body problem using Jacobi coordinates in section 1.2. In section 1.3, we derive the classical solutions for the two-body problem. In section 1.4 regularization is defined and collision is regularized via the classical Levi-Civita transformation and the Kustaanheimo-Stiefel transformation. Section 1.5 introduces the equations of motion for the restricted three-body problem in different variations and coordinate systems. This problem is the main focus of subsequent chapters. Also, we discuss briefly in section 1.1 the global behavior of solutions in the n-body problem having collision and noncollision singularities. Key results are stated, including Sundman's basic theorems and the Painlevé conjecture proven by Xia. This material serves as background introductory material and provides an historical perspective. The integrals of motion for the n-body problem are also derived. In section 1.6 geodesic equivalent flows on spaces of constant curvature are derived using the Euler-Lagrange differential equations, and their equivalence with the flow of the two-body problem is described. A new proof is given for this equivalence which is substantially shorter than previous proofs. The geodesic flows give rise to n-dimensional regularizations.

1.1 THE N-BODY PROBLEM

We consider the n-body problem, $n \geq 2$. Of particular interest will be the case $n = 3$ for the Newtonian three-body problem. Before this problem and variations of it are defined, we define the general n-body problem and discuss existence of solutions. The basic conservation laws are derived.

It is defined by the motion of $n \geq 2$ mass particles $\mathrm{P_k}$ of masses $m_k > 0, k = 1, 2, \ldots, n$, moving in three-dimensional space x_1, x_2, x_3 under the classical Newtonian inverse square gravitational force law. We assume the

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Cartesian coordinates of the kth particle are given by the real vector $\mathbf{x}_k = (x_{k1}, x_{k2}, x_{k3}) \in \mathbf{R^3}$. The differential equations defining the motion of the particles are given by

$$m_k \ddot{\mathbf{x}}_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{G m_j m_k}{r_{jk}^2} \frac{\mathbf{x}_j - \mathbf{x}_k}{r_{jk}}, \tag{1.1}$$

$k = 1, 2, \ldots, n$, where $r_{jk} = |\mathbf{x}_j - \mathbf{x}_k| = \sqrt{\sum_{i=1}^{3} (x_{ji} - x_{ki})^2}$ is the Euclidean distance between the kth and jth particles, and $\cdot \equiv \frac{d}{dt}$. Equation (1.1) expresses the fact that the acceleration of the kth particle P_k is due to the sum of the forces of the $n-1$ particles $P_i, i = 1, \ldots, n, i \neq k$. The time variable $t \in \mathbf{R^1}$. Equation (1.1) represents $3n$ second order differential equations. This equation can be put into a simpler form by first dividing both sides through by m_k, and expressing it as a first order system,

$$\dot{\mathbf{x}}_k = \mathbf{v}_k, \qquad \dot{\mathbf{v}}_k = m_k^{-1} \frac{\partial U}{\partial \mathbf{x}_k}, \tag{1.2}$$

where $\mathbf{v}_k = (v_{k1}, v_{k2}, v_{k3}) = (\dot{x}_{k1}, \dot{x}_{k2}, \dot{x}_{k3}) \in \mathbf{R^3}$ are the velocity vectors of the kth particle,

$$U = \sum_{\substack{j=1 \\ j \neq k}}^{n} \frac{G m_j m_k}{r_{jk}},$$

$U = U(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is a real-valued function of $3n$ variables $x_{kj}, j = 1, 2, 3$, and

$$\frac{\partial U}{\partial \mathbf{x}_k} \equiv \left(\frac{\partial U}{\partial x_{k1}}, \frac{\partial U}{\partial x_{k2}}, \frac{\partial U}{\partial x_{k3}} \right),$$

$k = 1, 2, \ldots, n$. Equation (1.2) represents a system of $6n$ first order differential equations for the $6n$ variables $x_{k\ell}, v_{k\ell}, k = 1, 2, \ldots, n; \ell = 1, 2, 3$. U is the *potential energy*. G is the universal gravitational constant.

If we assume $r_{jk} > 0$, then U is a well-defined function and is a *smooth function* in the $3n$ variables x_{jk}, where *smooth* means that U has continuous partial derivatives of all orders in the variables x_{jk} and is real analytic. For notation, we set $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. Then $\mathbf{x} \in \mathbf{R^{3n}}$. Similarly, $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \in \mathbf{R^{3n}}$. With this notation, $U = U(\mathbf{x})$.

System (1.2) is of the form

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}), \tag{1.3}$$

where $\mathbf{y} = (\mathbf{x}, \mathbf{v}) \in \mathbf{R^{6n}}$, and also where $\mathbf{f} = (\mathbf{v}, m_1^{-1} \partial U / \partial \mathbf{x}_1, \ldots, m_n^{-1} \partial U / \partial \mathbf{x}_n) \in \mathbf{R^{6n}}$. Thus, the standard existence and uniqueness theorems of ordinary differential equations can be applied to (1.3), and hence (1.2).

Since $\mathbf{f} = (f_1, \ldots, f_{6n})$ is a smooth vector function of $\mathbf{y}$, then these theorems guarantee that through any initial point $\mathbf{y}(t_0) = \mathbf{y}_0$ at initial time

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

t_0 there exists a locally unique solution for $|t - t_0| < \delta$, where δ is sufficiently small. This can be made more precise: If the real functions f_k satisfy $|f_k| < M, k = 1, 2, \ldots, 6n$, in a domain $|\mathbf{y} - \mathbf{y}_0| < p$, then

$$\delta = \frac{p}{(1 + 6n)M}$$

(see [204]).

A system of integrals exist for (1.1) which can be used to reduce the dimension of the $(6n+1)$-dimensional coordinate space $(\mathbf{x}, \mathbf{y}, t)$. An *integral* is a real-valued function of the $6n + 1$ variables x_{kj}, v_{kj}, t which is constant when evaluated along a solution of (1.1). Let $\mathbf{x}(t), \mathbf{v}(t)$ represent a solution of (1.1).

Definition 1.1 *A integral of (1.1) is a real-valued function* $I(\mathbf{x}, \mathbf{v}, t)$ *such that*

$$\frac{d}{dt} I\left(\mathbf{x}(t), \mathbf{v}(t), t\right) = 0,$$

where the solution $\mathbf{x}(t), \mathbf{v}(t)$ *is defined.*

This definition implies that $I = c =$ constant along the given solution. This defines a $6n$-dimensional *integral manifold*,

$$I^{-1}(0) = \{(\mathbf{x}, \mathbf{v}, t) \in \mathbf{R^{6n+1}} | I = c\},$$

on which the solutions will lie.

Thus, an integral constrains the motion of the mass particles and can be used to reduce the dimension of the space of $6n+1$ coordinates, $x_{k\ell}, v_{k\ell}, t, k = 1, 2, \ldots, n; \ell = 1, 2, 3$ by 1, by solving for one of the coordinates as a function of the $6n$ remaining coordinates, at least implicitly. For notation we refer to the $6n$-dimensional real space of coordinates $(\mathbf{x}, \mathbf{v}) \in \mathbf{R^{3n}} \times \mathbf{R^{3n}}$, as the *phase space*, and $(\mathbf{x}, \mathbf{v}, t) \in \mathbf{R^{3n}} \times \mathbf{R^{3n}} \times \mathbf{R^1}$ as the *extended phase space.*

When two or more integrals $I_1(\mathbf{x}, \mathbf{v}, t), I_2(\mathbf{x}, \mathbf{v}, t)$ exist for (1.1), they are called *independent* if the gradient vectors $\partial_{\mathbf{x},\mathbf{v},t} \equiv (\partial_{\mathbf{x}_1}, \ldots, \partial_{\mathbf{x}_n}, \partial_{\mathbf{v}_1}, \ldots, \partial_{\mathbf{v}_n}, \partial_t)$ of I_1 and I_2 are independent. This implies that the rank of the $2 \times (6n + 1)$ matrix

$$\frac{\partial(I_1, I_2)}{\partial(\mathbf{x}, \mathbf{v}, t)}$$

is in general 2.

Equation (1.1) has a set of 10 independent algebraic integrals. These are given by the three classical conservation laws of linear momentum, energy, and angular momentum. We will derive these now.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

First, we derive the conservation of linear momentum. To do this, we add up the right side of (1.1),

$$S = \sum_{k=1}^{n}\sum_{j=1}^{n} \frac{Gm_j m_k}{r_{jk}^2}\frac{\mathbf{x}_j - \mathbf{x}_k}{r_{jk}},$$

$j \neq k$. $S = 0$ is verified, since each term $\mathbf{x}_j - \mathbf{x}_k$ occurs with its negative, and mutual cancellations occur for all the terms. This implies

$$\sum_{k=1}^{n} m_k \ddot{\mathbf{x}}_k = 0.$$

Setting

$$M = \sum_{k=1}^{n} m_k, \qquad \boldsymbol{\rho} = M^{-1}\sum_{k=1}^{n} m_k \mathbf{x}_k,$$

where $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3) \in \mathbf{R^3}$ is the *center of mass* vector of the particles, then $\ddot{\boldsymbol{\rho}} = \mathbf{0}$. This yields

$$\boldsymbol{\rho} = \mathbf{c}_1 t + \mathbf{c}_2, \tag{1.4}$$

$|t| < \delta$, where $\mathbf{c}_1, \mathbf{c}_2$ yield six constants which are uniquely determined from the initial conditions $\mathbf{x}_k(t_0), \mathbf{v}_k(t_0)$. Equation (1.4) expresses the law of the conservation of linear momentum: *The center of mass moves uniformly in a straight line.*

The origin of the Cartesian coordinate system x_1, x_2, x_3 for the motion of P_k can be shifted to the center of mass by setting $\bar{x}_j = x_j - \rho_j$. This does not alter the form of (1.1) since $\ddot{\rho}_j = 0$, and we can replace $\mathbf{x}_j$ by $\bar{\mathbf{x}}_j$. Thus, without loss of generality, we can assume $\boldsymbol{\rho} = \mathbf{0}$, which implies

$$\sum_{k=1}^{n} m_k \mathbf{x}_k = \mathbf{0}, \tag{1.5}$$

and also by differentiation,

$$\sum_{k=1}^{n} m_k \mathbf{v}_k = \mathbf{0}. \tag{1.6}$$

It is verified that (1.5), (1.6) represent six independent algebraic integrals $I_k, k = 1, 2, \ldots, 6$.

Another independent algebraic integral is given by the conservation of energy H,

$$H = T - U, \tag{1.7}$$

where H is the total energy of the system of n particles, and

$$T = \frac{1}{2}\sum_{k=1}^{n} m_k |\mathbf{v}_k|^2 \tag{1.8}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

is the *kinetic energy.* Thus, H is the sum of the potential and kinetic energies. It is an integral since one verifies by direct computation that $\frac{d}{dt}(T - U) = 0$ using (1.2). The law of conservation of energy states that the *energy is constant along solutions.*

The remaining three integrals are given by the conservation of angular momentum. This is derived by forming the vector cross product $\mathbf{x}_k \times \ddot{\mathbf{x}}_k$ using (1.1) and summing over k, where it is verified that

$$\sum_{k=1}^{n} m_k(\mathbf{x}_k \times \ddot{\mathbf{x}}_k) = \sum_{k=1}^{n}\sum_{j=1}^{n} \frac{Gm_j m_k}{r_{jk}^3}\mathbf{x}_k \times \mathbf{x}_j = \mathbf{0}, \tag{1.9}$$

where $j \neq k$ and where we used the fact $\mathbf{x}_k \times \mathbf{x}_k = \mathbf{0}$. The double sum is zero since $\mathbf{x}_j \times \mathbf{x}_k = -\mathbf{x}_k \times \mathbf{x}_j$. Integrating the left-hand side of (1.9) yields

$$\sum_{k=1}^{n} m_k(\mathbf{x}_k \times \mathbf{v}_k) = \mathbf{c}, \tag{1.10}$$

where $\mathbf{c} = (c_1, c_2, c_3) \in \mathbf{R^3}$ is the vector constant of angular momentum. Equation (1.10) expresses the law of conservation of angular momentum.

The angular momentum can be viewed as a measure of the rotational motion of (1.1). This measure of the rotation is illustrated in an important theorem of Sundman.

Theorem 1.2 (Sundman) *If at time $t = t_1$ all the particles P_k collide at one point, then $\mathbf{c} = \mathbf{0}$.*

This is called *total collapse.* The fact $\mathbf{c} = \mathbf{0}$ means that the particles are able to all collapse to a single location. In a sense, this is enabled because with $\mathbf{c} = \mathbf{0}$, the rotation has been taken away from the motion of the particles. For the two-body problem for $n = 2$, collision between P_1, P_2, where $r_{12} = 0$, can occur only if $\mathbf{c} = \mathbf{0}$. Theorem 1.2 is not proven here. (See [204].)

We conclude our introduction to the n-body problem with a brief summary of the extension of a solution $\mathbf{x}(t), \mathbf{v}(t)$ of (1.2) which has initial values $\mathbf{x}(t_0) = \mathbf{x}_0, \mathbf{v}(t_0) = \mathbf{v}_0$ at $t = t_0$. We extend this general solution for $t > t_0$. Now, either the $6n$ coordinates remain smooth for all time $t > t_0$, or else there is a first time $t = t_1$ where there is a *singularity* for at least one of the coordinates, where all coordinates are smooth for $t_0 \leq t < t_1$. The extent to which the solution can be continued in t beyond t_1 depends on whether or not, during the course of the motion of the P_k, the right hand-side of (1.2) remains smooth. Let $r_{\min}(t) = \min\{r_{jk}(t)\}, j < k$. $r_{\min}(t)$ is the minimum of the $n(n-1)/2$ distances r_{jk}. It can be proved that if t_1 is finite, then $r_{\min}(t) \to 0$ as $t \to t_1$. This implies $U \to \infty$ as $t \to t_1$. (See [204].)

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

In this case we say that there is a singularity of the solution at $t = t_1$. Surprisingly, this does not necessarily imply that a collision between the particles has to take place. This is called a *noncollision singularity.* The particles can get very close to each other and move in a complicated way so that the potential increases without bound. This question is a subtle one and is not considered in this book, as it is not the focus. However, we briefly summarize some key results on the nature of the singularity if $r_{\min} \to 0$ as $t \to t_1$.

If $n = 2$, then as $t \to t_1$ a collision must occur between m_1 and m_2. From Theorem 1.2, the condition that $\mathbf{c} = \mathbf{0}$ implies that the particles m_1, m_2 lie along a line, and as $r_{12} \to 0$ the collision can be *regularized.* This means that the solution can be smoothly continued to $t \geq t_1$ by a change of coordinates and time t. This is carried out in detail in sections 1.4 and 1.6. This means physically that m_1, m_2 perform a smooth bounce at collision and their motion for $t \geq t_1$ falls back along the same line on which they collided. Since $\mathbf{c} = \mathbf{0}$, three dimensions can be eliminated from the six-dimensional phase space. This means that the set of all collision orbits has a lower dimension than that of the phase space. The fact they have a lower dimension means that the total volume they make up is actually a set of relative zero volume in the full phase space. For example, in the two-dimensional Euclidean plane, the volume is *area* and all one-dimensional curves have zero area. The generalized volume of the phase space we use is called *measure*, which will mean *Lebesgue measure* [52]. Thus, the set of collision orbits in the two-body problem is a set of measure zero in the phase space. There is a natural way to assign a measure μ to the phase space $(\mathbf{x_1}, \mathbf{x_2}, \dot{\mathbf{x}}_\mathbf{1}, \dot{\mathbf{x}}_\mathbf{2}) \in \mathbf{R^{12}}$ of the three-dimensional two-body problem by setting

$$d\mu = dx_{11} dx_{12} dx_{13} \ldots dx_{33} d\dot{x}_{11} d\dot{x}_{12} d\dot{x}_{13} \ldots d\dot{x}_{33}.$$

This defines a twelve-dimensional volume element and generalizes in the natural way to the n-body problem, $n > 2$.

When $n = 3$ the situation is much more complicated. Two cases are considered. The first case is when $\mathbf{c} \neq \mathbf{0}$. By Theorem 1.2, simultaneous collision between all three particles cannot occur, and *only* binary collisions can occur between $P_k, P_\ell, k < \ell, k, \ell = 1, 2, 3$. As is shown in section 1.4, the set of all possible orbits leading to binary collisions in the eighteen-dimensional phase space is 16. Thus, they comprise a set of lower dimension and are of measure zero in the phase space. All these collisions in three dimensions can also be regularized by a transformation of position, velocity, and time, as described in Section 1.4, due to Sundman. This regularization uses a uniform time variable λ for all the collisions. After this transformation is applied, any solution can be continued through binary collision. Since only binary collisions can occur, Sundman was able to prove that any solution of

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

the three-body problem can be extended for all time and also be explicitly represented as a series expansion [216].

Theorem 1.3 (Sundman) *Any solution of the general three-body problem with* $\mathbf{c} \neq \mathbf{0}$ *can be continued for all time and represented as a series expansion in the time variable* λ *that represents the entire motion.*

This theorem is of important historical significance since finding a way to explicitly express the solutions of the n-body problem was an outstanding problem for many years prior to that time. Theorem 1.3 does not actually solve the three-body problem since it does not describe the actual dynamics. Nevertheless, it does represent a milestone.

To underscore the importance of trying to solve the n-body problem, King Oscar II of Sweden and Norway had established a prize for this in the latter 19th century. The prize was for finding a series expansion for the coordinates of the n-body problem valid for all time. Although Sundman indeed solved this for $n = 3$ in 1913, he did not receive the prize. Instead it went to Poincaré much earlier–in 1889. Even though Poincaré did not solve the problem, he was given the prize due to the large impact his work had on the entire field of dynamics. For a detailed proof of Theorem 1.3, see [204].

When $\mathbf{c} = \mathbf{0}$, the total collapse of the three particles can occur. The ability to regularize collision is related to the uniformity of collision among the three particles. As we saw in the case of two-body collision, collision is uniform when the two particles perform a smooth bounce. The fact that they collide at $t = t_1$ means that $U \to \infty$ as $t \to t_1$, and conversely. A noncollision singularity between three particles would imply that near collision between the three particles a smooth regularized flow would not be possible to achieve in general. On the other hand, nonregularizability of collision does not imply the existence of a noncollision singularity. Nevertheless, nonregularizability of collision is a necessary condition for noncollision singularity states, and its existence plays an important role.

The question of whether or not triple collision was regularizable was solved by McGehee [151] for the collinear three-body problem. In this case all three mass points lie on a line. The phase space is six-dimensional since three position coordinates are needed, one for each mass point, and correspondingly three velocity coordinates. The method of McGehee's proof is to introduce a change of coordinates and time so that the triple collision state is transformed into a lower dimensional manifold. This surface then corresponds to the state when all three masses simultaneously collide and is called the *McGehee triple collision manifold*. The nonregularizability of triple collision is realized by the property that solutions approaching near

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

triple collision, and hence near to the triple collision manifold, are led to widely divergent paths for small changes in their orbits. In the collinear three-body problem it can be seen that as $U \to \infty, r_{\min} \to 0$ implies that the three particles do in fact collide. The proof of this for the general three-body problem was given by Painlevé in the late 19th century. Thus, in this case nonregularizability of collision does not imply the existence of a noncollision singularity.

Thus, if a noncollision singularity occurs, it would have to occur in at least the four-body problem. This is the Painlevé conjecture. More precisely, the conjecture states that *for $n > 3$, there exist solutions with noncollision singularities.* This was stated in 1895. [72].

Inspired by Painlevé, von Zeipel proved the following interesting theorem in 1908 [222].

Theorem 1.4 (von Zeipel) *If a noncollision singularity occurs in the n-body problem, $n > 3$, then there would exist a solution which would become unbounded in finite time.*

This would seem to violate the fact that the speed of light is the maximal velocity for particles of mass, however, in our case P_k are of zero dimension, and thus the distance between P_k can become arbitrarily small and the velocities arbitrarily high.

The solution of the Painlevé conjecture was solved by Xia for the spatial five-body problem and published in 1992 [227].

Theorem 1.5 (Xia) *There exist noncollision singularities in the spatial five-body problem.*

For an historical exposition of the Painlevé conjecture and related topics, see [72]. The existence of a noncollision singularity in the spatial four-body problem is an open problem.

The focus of this book will be on solutions which generally do not collide. In fact, solutions will be studied for the three-body problem which do not collide and where the motion is chaotic. We will study this near special families of periodic orbits in chapter 2, and with the process of capture defined in chapter 3.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

1.2 PLANAR THREE-BODY PROBLEM

In later sections we will consider the planar circular restricted three-body problem as one of our main models. We derive it by first considering the general planar three-body problem, using Jacobi coordinates [191, 219].

The planar three-body problem is obtained from (1.2) by setting $n = 3$ and assuming $\mathbf{x}_k \in \mathbf{R^2}, \mathbf{v}_k \in \mathbf{R^2}, k = 1, 2, 3$.

The center of mass $\boldsymbol{\rho} = \mathbf{0}$, in accordance with (1.5), where $\boldsymbol{\rho} = M^{-1} \sum_{k=1}^{3} m_k\mathbf{x}_k, M = m_1 + m_2 + m_3, \sum_{k=1}^{3} m_k\mathbf{x}_k = \mathbf{0}$. It is useful to write out (1.1),

$$\begin{aligned} m_1\ddot{\mathbf{x}}_1 &= \frac{Gm_1m_2}{r_{12}^3}(\mathbf{x}_2 - \mathbf{x}_1) + \frac{Gm_3m_1}{r_{13}^3}(\mathbf{x}_3 - \mathbf{x}_1), \\ m_2\ddot{\mathbf{x}}_2 &= \frac{Gm_1m_2}{r_{12}^3}(\mathbf{x}_1 - \mathbf{x}_2) + \frac{Gm_3m_2}{r_{23}^3}(\mathbf{x}_3 - \mathbf{x}_2), \\ m_3\ddot{\mathbf{x}}_3 &= \frac{Gm_1m_3}{r_{13}^3}(\mathbf{x}_1 - \mathbf{x}_3) + \frac{Gm_3m_2}{r_{23}^3}(\mathbf{x}_2 - \mathbf{x}_3). \end{aligned} \tag{1.11}$$

This is a system of six second order differential equations. With the constraint $\boldsymbol{\rho} = \mathbf{0}$ we can eliminate one of the vector variables $\mathbf{x}_i$, resulting in four second order differential equations. Equations (1.11) are transformed to *Jacobi coordinates*. Set

$$\mathbf{q} = \mathbf{x}_2 - \mathbf{x}_1, \qquad \mathbf{Q} = \mathbf{x}_3 - \boldsymbol{\beta},$$

where $\boldsymbol{\beta} = \nu^{-1}(m_1\mathbf{x}_1 + m_2\mathbf{x}_2), \nu = m_1 + m_2$, is the center of mass vector of the binary pair m_1, m_2. $\mathbf{q}$ is the relative vector of P_2 with respect to P_1, and $\mathbf{Q}$ is the vector from $\boldsymbol{\beta}$ to P_3. (See Figure 1.1.) $\mathbf{q}, \mathbf{Q}$ are called Jacobi coordinates.

The transformation of (1.11) to Jacobi coordinates is now carried out. Each term $\mathbf{x}_k - \mathbf{x}_j, k > j, j, k = 1, 2, 3$, is transformed. First, by definition $\mathbf{q} = \mathbf{x}_2 - \mathbf{x}_1$. Next, $\mathbf{x}_3 - \mathbf{x}_1 = \mathbf{Q} + m_2\nu^{-1}\mathbf{q}$. This follows since $\mathbf{x}_3 - \mathbf{x}_1 = \mathbf{Q} + \boldsymbol{\beta} - \mathbf{x}_1 = \mathbf{Q} + \nu^{-1}(m_1\mathbf{x}_1 + m_2\mathbf{x}_2) - \mathbf{x}_1 = \mathbf{Q} + m_2(\mathbf{x}_2 - \mathbf{x}_1)\nu^{-1} = \mathbf{Q} + m_2\nu^{-1}\mathbf{q}$. In a similar way, $\mathbf{x}_3 - \mathbf{x}_2 = \mathbf{Q} - m_1\nu^{-1}\mathbf{q}$. Substituting the expressions for $\mathbf{x}_k - \mathbf{x}_j$ into (1.11), dividing the first differential equation by m_1, the second by m_2, and subtracting these two differential equations yields

$$\ddot{\mathbf{q}} = -\frac{G\nu}{|\mathbf{q}|^3}\mathbf{q} + Gm_3\left[\frac{\mathbf{Q} - m_1\nu^{-1}\mathbf{q}}{r_{23}^3} - \frac{\mathbf{Q} + m_2\nu^{-1}\mathbf{q}}{r_{13}^3}\right], \tag{1.12}$$

where $r_{12} = |\mathbf{q}|, r_{13} = |\mathbf{Q} + m_2\nu^{-1}\mathbf{q}|, r_{23} = |\mathbf{Q} - m_1\nu^{-1}\mathbf{q}|$. To obtain the differential equation for $\mathbf{Q}$, we use the third equation of (1.11). We need for this another relationship for $\mathbf{Q} = \mathbf{x}_3 - \boldsymbol{\beta} = \mathbf{x}_3 - \nu^{-1}(m_1\mathbf{x}_1 + m_2\mathbf{x}_2) = \mathbf{x}_3 - \nu^{-1}(-m_3x_3) = (1 + \nu^{-1}m_3)\mathbf{x}_3 = \nu^{-1}M\mathbf{x}_3$. This implies $\mathbf{x}_3 = \nu M^{-1}\mathbf{Q}$. Multiplying the third equation of (1.11) by $m_3^{-1}\nu^{-1}M$ yields

$$\ddot{\mathbf{Q}} = -\frac{GMm_1\nu^{-1}}{r_{13}^3}(\mathbf{Q} + m_2\nu^{-1}\mathbf{q}) - \frac{GMm_2\nu^{-1}}{r_{23}^3}(\mathbf{Q} - m_1\nu^{-1}\mathbf{q}). \tag{1.13}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

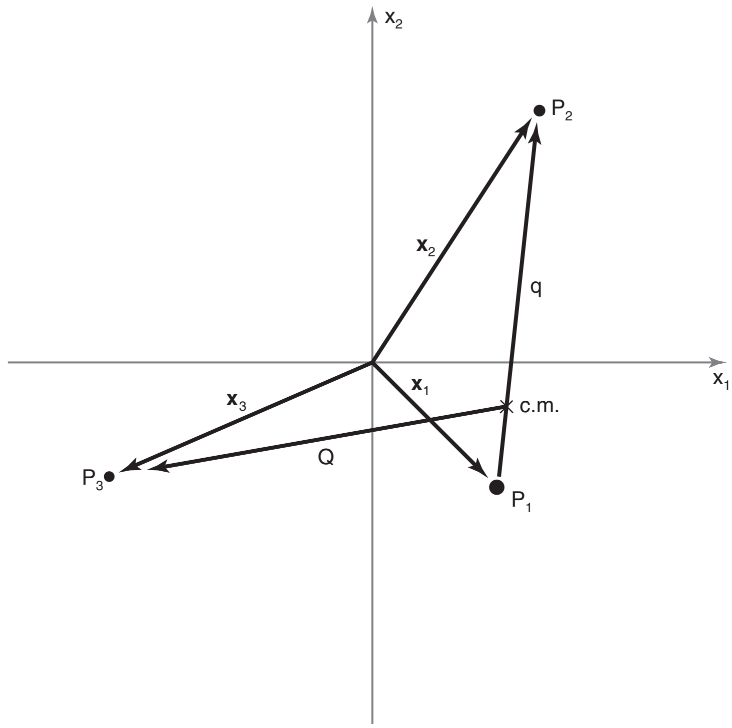

Figure 1.1 Jacobi coordinates.

Equations (1.12), (1.13) represent (1.11) in Jacobi coordinates, which are four second order differential equations for $(\mathbf{q}, \mathbf{Q}) = (q_1, q_2, Q_1, Q_2) \in \mathbf{R^4}$.

The kinetic energy T of this system given by (1.8) for $n = 3$ takes a nice form if we set

$$\dot{\mathbf{q}} = k_2^{-1}\mathbf{p}, \qquad \dot{\mathbf{Q}} = k_1^{-1}\mathbf{P}, \tag{1.14}$$

where

$$k_1 = m_3 \nu M^{-1}, \qquad k_2 = m_1 m_2 \nu^{-1}.$$

It is noted that if $m_3 = 0$, then $\mathbf{P} = \mathbf{0}$.

Lemma 1.6 *The kinetic energy of the system (1.12), (1.13) is given by*

$$T = \frac{1}{2}(k_2^{-1}|\mathbf{p}|^2 + k_1^{-1}|\mathbf{P}|^2) \tag{1.15}$$

and the total energy $\mathcal{H}$ of the system is

$$\mathcal{H} = T - U,$$
$$U = \frac{Gm_1m_2}{|\mathbf{q}|} + \frac{Gm_2m_3}{r_{23}} + \frac{Gm_1m_3}{r_{13}}. \tag{1.16}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Proof. The form of $\mathcal{H}$ follows from (1.15) together with (1.7). To transform (1.8), we use (1.14). It is verified that from $\mathbf{x}_3 - \mathbf{x}_2 = \mathbf{Q} - m_1\nu^{-1}\mathbf{q}$, $\mathbf{x}_3 = M^{-1}\nu\mathbf{Q}$,

$$\mathbf{x}_1 = -m_3 M^{-1}\mathbf{Q} - m_2\nu^{-1}\mathbf{q},$$
$$\mathbf{x}_2 = -m_3 M^{-1}\mathbf{Q} + m_1\nu^{-1}\mathbf{q}.$$

Then, (1.14) implies

$$\dot{\mathbf{x}}_1 = -\nu^{-1}\mathbf{P} - m_1^{-1}\mathbf{p},$$
$$\dot{\mathbf{x}}_2 = -\nu^{-1}\mathbf{P} + m_2^{-1}\mathbf{p},$$
$$\dot{\mathbf{x}}_3 = m_3^{-1}\mathbf{P}.$$

Substituting these relationships into (1.8) yields (1.15) after simplification. □

It is immediately verified that (1.12), (1.13) can be written in Hamiltonian form,

$$\dot{\mathbf{q}} = \mathcal{H}_{\mathbf{p}}, \quad \dot{\mathbf{p}} = -\mathcal{H}_{\mathbf{q}},$$
$$\dot{\mathbf{Q}} = \mathcal{H}_{\mathbf{P}}, \quad \mathbf{P} = -\mathcal{H}_{\mathbf{Q}}, \tag{1.17}$$

where $\mathcal{H}_{\mathbf{p}} \equiv \partial\mathcal{H}/\partial\mathbf{p}$.

Jacobi coordinates are particularly well suited to studying versions of the three-body problem where P_1, P_2 are performing a given binary motion and the mass of P_3 is infinitesimally small. This situation would occur, for example, if P_3 were considered to be a small object such as a spacecraft, comet, or asteroid, if P_1 were the Sun, and if P_2 were Jupiter. For all practical purposes, m_3 has negligible mass. Setting $m_3 = 0$ reduces (1.12) to

$$\ddot{\mathbf{q}} = \frac{G\nu}{|\mathbf{q}|^3}\mathbf{q}, \tag{1.18}$$

which represents the standard differential equation for two-body problem for the motion of P_1, P_2 in relative coordinates centered at P_1. Equation (1.18) can be explicitly solved. Equation (1.13) then describes the motion of P_3 in the gravitational field generated by the motion of the particles P_1, P_2 defined by (1.18).

Before proceeding to the restricted three-body problem, the solutions of (1.18) are derived.

1.3 TWO-BODY PROBLEM

The *Keplerian two-body problem* is defined by (1.18) and is in the coordinates shown in Figure 1.2.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

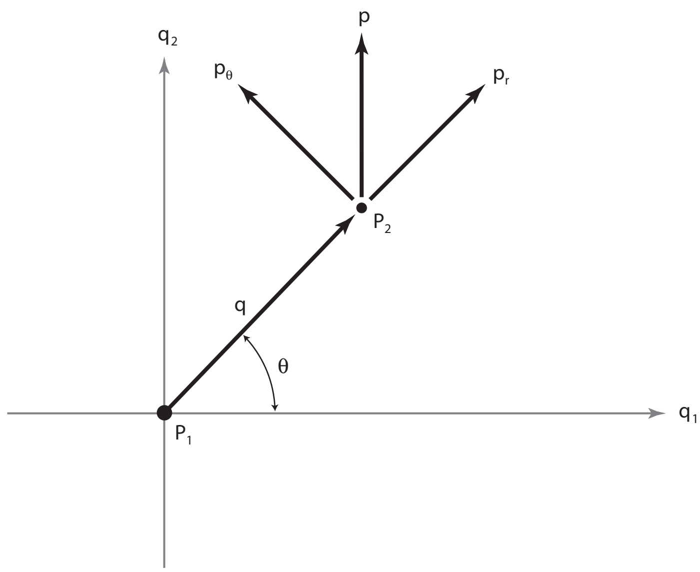

Figure 1.2 Polar coordinates and momenta.

The total energy of the system is given by the real-valued function, $H(p, q) : \mathbf{R^4} \to \mathbf{R^1}$,

$$H = \frac{1}{2}|\mathbf{p}|^2 - \frac{G\nu}{|\mathbf{q}|}, \tag{1.19}$$

where $\mathbf{p} = \dot{\mathbf{q}} \in \mathbf{R^2}$. It is noted that the coefficient of $\frac{1}{2}|\mathbf{p}|^2$ is 1 and does not contain m_1, m_2 as in the general form of the energy (1.7). This is because (1.7) is written in terms of inertial coordinates, and (1.18) is in relative coordinates. $\mathbf{p}$ are referred to as the *momentum* variables, which can be viewed as linear momentum of a unit mass since $\mathbf{p} = 1\dot{\mathbf{q}}$. Equation (1.18) is the first order system

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = -\frac{G\nu}{|\mathbf{q}|^3}\mathbf{q}, \tag{1.20}$$

or, in Hamiltonian form, using (1.19),

$$\dot{\mathbf{q}} = H_{\mathbf{p}}, \quad \dot{\mathbf{p}} = -H_{\mathbf{q}}. \tag{1.21}$$

H is called the *Hamiltonian function*. The solutions of (1.20) are well known and depend on the value $H = h$ of the energy. If $h < 0$, the curve $\mathbf{q}(t)$ is an ellipse with a focus at P_1. For $h = 0, \mathbf{q}(t)$ is a parabola, and for $h > 0, \mathbf{q}(t)$ is a hyperbola, where both the parabola and hyperbola have foci at the origin, or alternatively P_1. These solutions are derived in this section and section 1.6.

The angular momentum is given by

$$\mathbf{c} = \mathbf{q}\times\mathbf{p} = q_1p_2 - q_2p_1. \tag{1.22}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

We prove that

$$|\mathbf{c}| = c = r^2\dot{\theta}, \tag{1.23}$$

$r = |\mathbf{q}|, \theta$ is the polar angle shown in Figure 1.2. Referring to Figure 1.2, $\mathbf{p}$ can be decomposed into its tangential and radial components, $\mathbf{p}_\theta, \mathbf{p_r}$, respectively, $\mathbf{p} = \mathbf{p}_r + \mathbf{p}_\theta = p_r\mathbf{e}_r + p_\theta\mathbf{e}_\theta$, where $\mathbf{e}_r, \mathbf{e}_\theta$ are unit vectors in the r, θ directions, respectively.

Differentiating $\mathbf{q} = r\mathbf{e}_r$ implies $\dot{\mathbf{q}} = \mathbf{p} = r\dot{\mathbf{e}}_r + \dot{r}\mathbf{e}_r$, where $\dot{\mathbf{e}}_r = \dot{\theta}\mathbf{e}_\theta$, which follows since $\mathbf{e}_r$ rotates with constant circular velocity $\dot{\theta}$. (Similarly, $\dot{\mathbf{e}}_\theta = -\dot{\theta}\mathbf{e}_r$.) Thus, $\mathbf{c} = \mathbf{q}\times\mathbf{p} = r^2\dot{\theta}\mathbf{e}_r\times\mathbf{e}_\theta$, verifying (1.23). The constancy of $\boldsymbol{c}$ gives the law discovered by Kepler that as m_2 moves in an elliptic orbit about m_1, it traces out equal areas in equal times. This is *Kepler's first law*. This follows since the change of area $A(t)$ that is swept out by m_2 in time t is approximated by $\Delta A = \frac{1}{2}r^2\dot{\theta}\Delta t$ for $\Delta t \ll 1$, since the base of the triangle in Figure 1.3 has length $rd\theta \approx r\dot{\theta}\Delta t$.

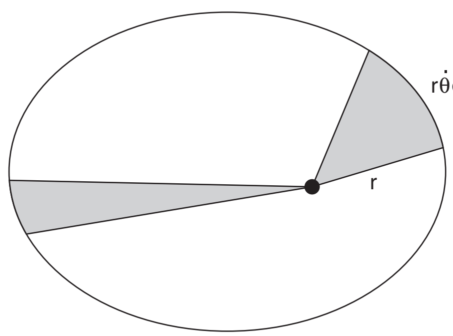

Figure 1.3 Sectorial area.

Therefore, as $\Delta t \to 0, \dot{A} = \frac{1}{2}c =$ constant. $\dot{A}$ is called the *sectorial area*.

The constancy of $\mathbf{c}$ also implies that all orbits are planar. Letting $\mathbf{q} \in \mathbf{R^3}, \mathbf{p} \in \mathbf{R^3}$, then $\mathbf{c}\times\mathbf{q} = (\mathbf{q}\times\mathbf{p})\times\mathbf{q} = \mathbf{0}$, which means that $\mathbf{q}(t)$ is perpendicular to $\mathbf{c}, t \in \mathbf{R^1}$. Thus, in (1.18) it is sufficient to assume $\mathbf{q} \in \mathbf{R^2}$.

Using a special regularizing transformation defined in the following sections, the solutions of (1.21) are explicitly determined and are nonsingular in the entire phase space. However, we derive here for future reference the solution for (1.18) or (1.21) with $h < 0$. Equation (1.18) is solved by transforming it to polar coordinates r, θ. Differentiating the previously stated polar representation for $\dot{\mathbf{q}}$ yields

$$\ddot{\mathbf{q}} = (\ddot{r} - r\dot{\theta}^2)\mathbf{e}_r + (2\dot{r}\dot{\theta} + r\ddot{\theta})\mathbf{e}_\theta.$$

On the other hand, (1.18) implies $\ddot{\mathbf{q}} = U_\mathbf{q}, U = G\nu/r$ and since $U_\mathbf{q}$ represents a central force field which is radially directed,

$$U_\mathbf{q} = U_r\mathbf{e}_r.$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Thus, equating coefficients in $\ddot{\mathbf{q}}$ yields

$$\ddot{r} - r\dot{\theta}^2 = U_r, \qquad 2\dot{r}\dot{\theta} + r\ddot{\theta} = 0.$$

Since $r^2\dot{\theta} = c =$ constant, then knowing $r(t)$ we can determine $\dot{\theta}(t)$, and hence $\theta(t)$ by quadratures. Therefore, it is sufficient to solve the first differential equation for $r(t)$, which we can write as

$$\ddot{r} = V_r, \tag{1.24}$$

where

$$V = U - \frac{c^2}{2r^2}$$

is called the *effective potential energy.* Thus, we have proved

Lemma 1.7 *Equation (1.18) can be reduced to solving (1.24).*

Lemma 1.8 *The total energy associated with (1.24) is given by*

$$H = \frac{1}{2}\dot{r}^2 - V. \tag{1.25}$$

Proof. In polar coordinates, (1.19) becomes

$$H = \frac{1}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - U = \frac{1}{2}(\dot{r}^2 + r^2\dot{\theta}^2) - V - \frac{c^2}{2r^2},$$

and $c = r^2\dot{\theta}$ yields (1.25). □

Since H is a real constant, then (1.25) can be used to solve (1.24) by solving for $\dot{r}$, and using quadrature,

$$\int dt = \int \frac{dr}{\sqrt{2(H - V(r))}},$$

which implicitly yields $r = r(t)$.

This integral equation is used to solve explicitly for $r = r(\theta)$. Using this expression and the one for c, we obtain

$$\frac{d\theta}{dt} = \frac{d\theta}{dr}\frac{dr}{dt} = \frac{d\theta}{dr}\sqrt{2(H - V(r))}.$$

Thus

$$\theta = \int \frac{(c/r^2)dr}{\sqrt{2(H - V(r))}}. \tag{1.26}$$

θ is called the *true anomaly.*

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Lemma 1.9 *Equation (1.26) implies*

$$r = \frac{p}{1 + e\cos\theta}, \tag{1.27}$$

where

$$p = c^2/k, \quad e = \left(1 + \frac{2Hc^2}{k^2}\right)^{\frac{1}{2}}, \quad k = G\nu.$$

Proof. Integration of (1.26) gives

$$\theta(r) = \cos^{-1}\left[\frac{cr^{-1} - c^{-1}k}{\left(2H + \frac{k^2}{c^2}\right)^{\frac{1}{2}}}\right] + \theta(r_0)$$

which gives (1.27), where $\theta(r_0) = 0$ and $r_0 = r_p, r_p$ is the periapsis radius, which is the radius of closest approach. □

Equation (1.27) represents an ellipse, where $H = h < 0, e < 1$. e is the *eccentricity*, and p is the *semilatus rectum*. (See Figure 1.4.) The focus of the ellipse is at the origin, where m_1 is located. This is *Kepler's second law*, that the attracting mass occupies a focus of the particle's orbit. r_a is the furthest point on the ellipse to the origin, called the *apoapsis*.

Different relationships between the parameters a, e, p, etc. can be derived, where a is the *semimajor axis*. For example, $g = ae$ is the distance from the center of the ellipse to the focus. This implies $f^2 = a^2 - g^2 = a^2(1 - e^2)$. Also, $a = p/(1 - e^2)$. It is verified that $r_p = a(1 - e), r_a = a(1 + e)$. For a

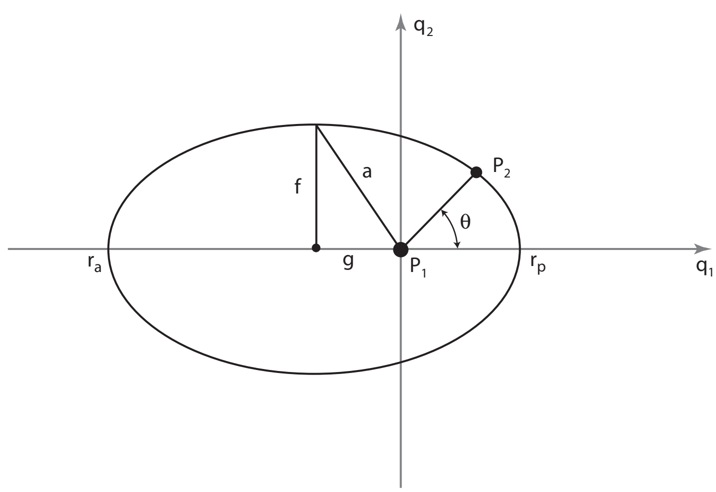

Figure 1.4 Geometry of the Kepler ellipse.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

circular orbit, it is checked that $e = 0$, which implies $r = p = a =$ constant. Equation (1.27) is often expressed as

$$r = \frac{a(1-e^2)}{1+e\cos\theta}. \tag{1.28}$$

An important relationship is the connection between H and a. From $e = (1 + 2Hc^2/k^2)^{\frac{1}{2}}, a = p/(1-e^2)$, we obtain

$$H = -\frac{k}{2a}. \tag{1.29}$$

Thus, a is a measure of the amount of energy contained in the Kepler motion.

Kepler's third law says that the square of the period of motion, P, along the ellipse (1.27) is proportional to the cube of a. This is derived using $A(t)$. Since $\frac{dA}{dt} = \frac{1}{2}c$, then

$$A(P) = \frac{1}{2}cP = \text{area of the ellipse} = \pi a f.$$

Therefore, since $f = a\sqrt{1-e^2}$,

$$P = 2\pi c^{-1}a^2\sqrt{1-e^2},$$

which by substitution of (1.29) and $e = (1 + 2Hc^2/k^2)^{\frac{1}{2}}$, yields

$$P^2 = (2\pi)^2 k^{-1} a^3, \tag{1.30}$$

which is Kepler's third law.

Finally, the *mean anomaly* is given by an angle $\mathcal{M}$ varying during one revolution from 0 to 2π with constant angular velocity. Since $k^{\frac{1}{2}}P/a^{\frac{3}{2}} = 2\pi$, then

$$\mathcal{M} = \frac{k^{\frac{1}{2}}}{a^{\frac{3}{2}}}t.$$

The *eccentric anomaly* E is defined as

$$\mathcal{M} = E - e\sin E.$$

It is seen in the next section that E is desirable to use in place of t to describe the Kepler motion.

When $H \geq 0$, an analogous description of the parabolic and hyperbolic motion can be made (see [214]). In the next section all values of H are treated in a uniform manner.

1.4 REGULARIZATION OF COLLISION

Equation (1.20) is singular when $\mathbf{q} = \mathbf{0}$, where, as $\mathbf{q} \to \mathbf{0}, U = k|\mathbf{q}|^{-1} \to \infty$. A solution $\mathbf{q}(t), \mathbf{p}(t)$ of (1.20), where $\mathbf{q}(t) \to \mathbf{0}, |\mathbf{p}(t)| \to \infty$ as $t \to t_0$, is

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

called a *collision orbit*, where collision is given by $\mathbf{q} = 0$ at $t = t_0$ between m_1 and m_2. Let $\boldsymbol{\varphi}(t) = (\mathbf{q}(t), \mathbf{p}(t)) \in \mathbf{R^4}$ represent a collision orbit. We show how to continue a solution through collision in a smooth fashion.

Let

$$\mathbf{q} = \mathbf{q}(\mathbf{u}, \mathbf{w}), \qquad \mathbf{p} = \mathbf{p}(\mathbf{u}, \mathbf{w}), \qquad t = t(\mathbf{u}, \mathbf{w}, s) \tag{1.31}$$

be a transformation, where $\mathbf{u} \in \mathbf{R^m}, \mathbf{w} \in \mathbf{R^m}, s \in \mathbf{R^1}, m \geq 2$, which transforms (1.20) into a system

$$\mathbf{u}' = \mathbf{f}(\mathbf{u}, \mathbf{w}), \qquad \mathbf{w}' = \mathbf{g}(\mathbf{u}, \mathbf{w}), \tag{1.32}$$

$' \equiv \frac{d}{ds}$, and which transforms $\boldsymbol{\varphi}(t)$ into $\boldsymbol{\zeta}(s) = (\mathbf{u}(s), \mathbf{w}(s))$. We assume that $s = s_0 < \infty$ corresponds to collision, where $\mathbf{q} = 0$ corresponds to $\mathbf{u} = \mathbf{u}_0$ and $\mathbf{w} \in \mathcal{C}, \mathcal{C}$ is a bounded subset of $\mathbf{R^m}$. $\mathbf{u}$ is the transformed position vector, and $\mathbf{w}$ is the transformed velocity vector. Assume that a given collision orbit $\boldsymbol{\varphi}(t), t \neq t_0$, is mapped into $\boldsymbol{\zeta}(s), s \neq s_0$, and as $s \to s_0, \mathbf{w}(s) \to \mathbf{w}_0 \in \mathcal{C}$.

Definition 1.10 *If (1.32) is smooth in a neighborhood of the set* $A = \{(\mathbf{u}_0, \mathbf{w}), \mathbf{w} \in \mathcal{C}\}$, *which are assumed to be nonequilibrium points of (1.32), then (1.20) is* regularizable *at* $\mathbf{q} = 0$, *and (1.31) is called a* regularization.

Thus, by this definition any collision solution $\boldsymbol{\zeta}(s)$ can be extended smoothly as a function of s through $s = s_0$, for s in a neighborhood of s_0. Also, any collision solution $\boldsymbol{\zeta}(s, \tilde{\mathbf{u}}_\mathbf{0}, \tilde{\mathbf{w}}_\mathbf{0})$ is a smooth function of $\tilde{\mathbf{u}}_\mathbf{0}, \tilde{\mathbf{w}}_\mathbf{0}$, where $\tilde{\mathbf{u}}_\mathbf{0}, \tilde{\mathbf{w}}_\mathbf{0}$ are in a neighborhood of A. In this case (1.31) is called a *local regularization*. If (1.32) is smooth for all $(\mathbf{u}, \mathbf{w}) \in \mathbf{R^4}$, then (1.31) is called a *global regularization*. This is not the most general definition of regularization, but it is sufficient for our presentation.

We consider a special regularization due to Levi-Civita [134]. In complex notation it is given by

$$q = z^2, \quad p = \frac{wz}{2|z|^2}, \tag{1.33}$$

where $q = q_1 + iq_2 \in \mathbb{C}, w = w_1 + iw_2 \in \mathbb{C}, i^2 = -1, |z|^2 = z\bar{z}, \bar{z} = z_1 - iz_2$. Thus $q = 0$ is mapped into $z = 0$.

(1.33) is written in *canonical form*. This means that it preserves the differential form $w = dp_1dq_1 + dp_2dq_2 = Re(\overline{dp}dq) = Re(\overline{dw}dz)$, where $Re(z) = z_1$.

Lemma 1.11 ω *is invariant under (1.33).*

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Proof. We show that (1.33) satisfies $\bar{p}dq = \bar{w}dz$ under (1.33). Lemma 1.11 follows by applying the operator d to both sides of $\bar{p}dq = \bar{w}dz$ and noting that $d^2 = 0$. Assume $q = z^2$. Then

$$\bar{p}dq = 2\bar{p}zdz.$$

Thus, setting $\bar{w} = 2\bar{p}z$ yields the invariance of ω and implies

$$p = \frac{w}{2\bar{z}} = \frac{wz}{2|z|^2},$$

giving (1.33). □

The invariance of ω implies that the area elements of the phase space are preserved. By Liouville's theorem, this implies that the Hamiltonian form of the differential equations are preserved [14]. Thus, the form of (1.21) is preserved in the new coordinates. If

$$\Phi(\mathbf{z}, \mathbf{w}) = H(\mathbf{q}(\mathbf{z}), \mathbf{p}(\mathbf{z}, \mathbf{w})),$$

then

$$\dot{\mathbf{z}} = \Phi_{\mathbf{w}},\ \mathbf{w} = -\Phi_{\mathbf{z}}, \tag{1.34}$$

$\mathbf{z} = (z_1, z_2) \in \mathbf{R^2}, \mathbf{w} = (w_1, w_2) \in \mathbf{R^2}$, and

$$\Phi(\mathbf{z}, \mathbf{w}) = \frac{1}{8}|\mathbf{z}|^{-2}|\mathbf{w}|^2 - k|\mathbf{z}|^{-2}. \tag{1.35}$$

We restrict H to the value $H = h$, so that (1.21) is defined on the three-dimensional energy level $H^{-1}(h) = \{(\mathbf{p}, \mathbf{q}) \in \mathbf{R^4} | H(\mathbf{p}, \mathbf{q}) = h\}$. Thus, (1.34) is defined on the set

$$\Phi^{-1}(h) = \{(\mathbf{z}, \mathbf{w}) \in \mathbf{R^4} | \Phi(\mathbf{z}, \mathbf{w}) = h\}.$$

The Levi-Civita transformation is augmented by the time transformation

$$dt = |\mathbf{q}|ds = |\mathbf{z}|^2 ds, \tag{1.36}$$

which as $|\mathbf{z}| \to 0$ stretches the new time variable s. This implies that (1.34) takes the form

$$\mathbf{z}' = |\mathbf{z}|^2\Phi_{\mathbf{w}}, \qquad \mathbf{w}' = -|\mathbf{z}|^2\Phi_{\mathbf{z}},$$

$' \equiv \frac{d}{ds}$, or equivalently

$$\mathbf{z}' = \tilde{\Phi}_{\mathbf{w}}, \qquad \mathbf{w}' = -\tilde{\Phi}_{\mathbf{z}}, \tag{1.37}$$

where $\tilde{\Phi} = (\Phi - h)|\mathbf{z}|^2$, or

$$\tilde{\Phi} = \frac{1}{8}|\mathbf{w}|^2 - h|\mathbf{z}|^2 - k \tag{1.38}$$

and $\mathbf{w}, \mathbf{z}$ are restricted to the set where $\tilde{\Phi} = 0$.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Theorem 1.12 *The Kepler problem given by (1.21) on the three-dimensional energy surface* $H^{-1}(h), h \in \mathbf{R^1}$*, is transformed by the Levi-Civita transformation (1.33) together with the time transformation (1.36) into the second order linear system*

$$\mathbf{z}'' - \frac{h}{2}\mathbf{z} = 0. \tag{1.39}$$

Collision for (1.21) is mapped onto the set $A = \{\mathbf{z}, \mathbf{w} | \mathbf{z} = 0, |\mathbf{w}|^2 = 8k\}, k = G\nu \neq 0$*, where (1.39) is smooth and* $\mathbf{w} = \mathbf{4z'}$.

Proof. Equation (1.37) implies

$$\mathbf{z}' = \frac{1}{4}\mathbf{w}, \qquad \mathbf{w}' = 2h\mathbf{z},$$

which yields (1.39). Setting $\mathbf{z} = 0$ in (1.38) yields $|\mathbf{w}|^2 = 8k$, and hence the set A. Equation (1.39) is smooth at $\mathbf{z} = 0$. □

Lemma 1.13 *Equations (1.33), (1.36) represent a global regularization of (1.21).*

Proof. A does not represent a location where (1.39) has equilibrium points. Thus, any collision solution can be extended up to $\mathbf{z} = 0$ at a finite time $t = t_0 < \infty$, and then extended beyond collision for $t > t_0$. Thus, (1.33), (1.36) is a local regularization. It is a global regularization since (1.39) is smooth on the entire energy surface $\tilde{\Phi}^{-1}(0)$. □

In [23], examples of local regularizations are given. Ignoring canonical extensions and the time transformation, an example of one for (1.35) in the momentum coordinates is given by the so-called *Sundman transformation*,

$$\mathbf{p} \rightarrow \frac{\mathbf{p}}{|\mathbf{p}|^2} = \tilde{\mathbf{p}}.$$

$|\mathbf{p}| = \infty$ is mapped into $\tilde{\mathbf{p}} = \mathbf{0}$. Thus $|\mathbf{p}| = \infty$ is mapped into a finite point where a local regularization can be constructed by canonically extending this map to a map of the $\mathbf{q}$ variables and t. However, this is not a global regularization, since in the new coordinates $\tilde{\mathbf{p}}, \tilde{\mathbf{q}}, |\tilde{\mathbf{p}}| = \infty$ will be a singular point and it corresponds to the point $\mathbf{p} = \mathbf{0}$, which is smooth for (1.21). Thus, the Sundman transformation introduces another singularity into the phase space $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) \in \mathbf{R^4}$.

It is remarked that A in Theorem 1.12 is equivalent to the set $\{\mathbf{z} = 0\} \times S^1$, where S^1 is a circle of radius $\sqrt{8k}$. Thus, collision for (1.21) has been reduced to the set product of a point and a circle.

It is noted that $c = 0$ is required for collision to occur. For the case of $h < 0$ corresponding to elliptic motion, for example, this implies that $e = 1$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

by (1.27) for any solution leading to collision. By (1.29), for any finite $h < 0$, this implies that a is finite, and so is r_a, where $r_p = 0$. Thus as $e \uparrow 1$, or equivalently as $c \to 0$, a family of ellipses are obtained which get thinner and thinner and converge to a line representing a collision orbit, where $\theta =$ constant as follows from c. The existence of a smooth regularization means that m_2 collides with m_1 at a finite time, smoothly bounces off of m_1, moves away from m_1 along the same line, and then repeats the process. This is called a *consecutive collision orbit* [23, 219]. (See Figure 1.5.)

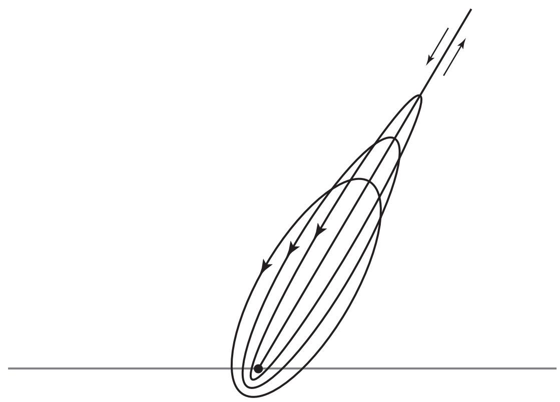

Figure 1.5 Evolution of ellipses to a consecutive collision orbit.

Lemma 1.14 *The set of collision orbits for (1.21) on each three-dimensional energy surface $H^{-1}(h)$ is topologically equivalent to the two-dimensional set $S^1 \times \mathbf{R^1}$.*

Proof. By Theorem 1.12, the collision manifold corresponds to the circle S^1, where $\mathbf{z} = \mathbf{0}$. Each point $\mathbf{w}_0$ on this circle can therefore be viewed as an initial value on the collision manifold of a collision orbit for $s = s_o$. Since the Levi-Civita regularization is global, each collision orbit is defined for all time $s \in \mathbf{R^1}$. Thus, the set of all collision orbits corresponds to the cylinder $C^2 = S^1 \times \mathbf{R^1}$. □

By Lemma 1.14, the set of all collision orbits in the four-dimensional phase space is then obtained by varying $h \in \mathbf{R^1}$, so we obtain the three-dimensional set $C^2 \times \mathbf{R^1}$. The dimension of this set agrees with the dimension of the set $\{\mathbf{c} = \mathbf{0}\}$.

Lemma 1.15 *The set of all collision orbits in the general planar three-body problem for $\mathbf{c} \neq \mathbf{0}$ is a set of dimension 11 in the 12-dimensional phase space.*

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Proof. From the preceding discussion following the proof of Lemma 1.14, the set of all collision orbits in the planar two-body problem is a set of three dimensions in the four-dimensional phase space. Assume a binary collision occurs at time $t = t_0$ for (1.2), $\mathbf{x}_k \in \mathbf{R^2}, k = 1, 2, 3$. $t_0 < \infty$ is well defined since $\mathbf{c} \neq \mathbf{0}$ where only isolated binary collisions can occur. The collision manifold for any of the three possible binary collisions is then three dimensions. On the other hand, the center of mass of the binary collision can no longer be fixed, as conservation of linear momentum is not valid due to the presence of a third mass point. Thus, the binary collision depends on the location in phase space of the two masses relative to the third. This introduces eight free parameters due to the position and velocity of two mass points. This, together with the three dimensions of the collision manifold, yields eleven dimensions, in the twelve dimensional phase space. □

It is noted that in the proofs of Lemmas 1.14 and 1.15, the Levi-Civita regularization need not be used to deduce the dimension of the collision manifold and set of all collision orbits. Any regularization can be used locally or globally. In [216] a local regularization is in fact used in the case of the three-dimensional three-body problem. Consider the three-dimensional two-body problem defined by (1.20) with $\mathbf{p} \in \mathbf{R^3}, \mathbf{q} \in \mathbf{R^3}$. In that case it is verified that the collision manifold in the six-dimensional phase space for (1.20) is three dimensions, and the set of all collision orbits is then four dimensions where time is the fourth dimension. When another mass is included and we have a three-body problem in three dimensions given by (1.2) with $n = 3, \mathbf{x}_k \in \mathbf{R^3}, k = 1, 2, 3$, then the phase space is eighteen dimensional, and for any binary collision, twelve additional free parameters are required. Thus, in this higher dimensional case, the set of all binary collision orbits in the general three-body problem in three dimensions with $\mathbf{c} \neq 0$ is therefore sixteen dimensions.

Thus, in the planar or three-dimensional general three-body problem with $\mathbf{c} \neq \mathbf{0}$, the set of all binary collisions is of smaller dimension than the dimension of the phase space, and hence a set of measure zero. This is summarized in the following lemma.

Lemma 1.16 *The set of all binary collision orbits in the general three-dimensional three-body problem with* $\mathbf{c} \neq 0$ *is a set of measure zero.*

The set of all solutions of (1.39) can be explicitly determined.

For $h < 0$, we write (1.39) as $\mathbf{z}'' + \frac{|h|}{2}\mathbf{z} = 0$. This is just a harmonic oscillator whose general solution is given by

$$\mathbf{z}(s) = \mathbf{a}_1 \cos \lambda s + \mathbf{a}_2 \sin \lambda s, \tag{1.40}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

$\mathbf{a}_i = (a_{i1}, a_{i2}) \in \mathbf{R^2}, i = 1, 2, \lambda = \sqrt{\frac{|h|}{2}}$.

When $h = 0$, (1.39) reduces to $\mathbf{z}'' = 0$, which yields

$$\mathbf{z}(s) = \mathbf{a}_1 s + \mathbf{a}_2, \tag{1.41}$$

and for $h > 0$,

$$\mathbf{z}(s) = \mathbf{a}_1 e^{\lambda s} + \mathbf{a}_2 e^{-\lambda s}, \tag{1.42}$$

$\lambda = \sqrt{\frac{h}{2}}$.

In section 1.6 the case of general energy is studied from a differential geometric viewpoint.

We conclude this section by considering the case $h < 0$. We show that the position map associated with the Levi-Civita transformation, written in complex coordinates,

$$q = z^2, \tag{1.43}$$

$q \in \mathbb{C}, z \in \mathbb{C}$, maps elliptic motion into elliptic motion. Set $q(\theta) = r(\theta)e^{i\theta}$, where $r = r(\theta)$ is given by (1.27). Equation (1.43) implies

$$z = r^{\frac{1}{2}} e^{i\frac{\theta}{2}},$$

which again represents an elliptical orbit, $z = z(\theta)$. More precisely, in component form, (1.43) is

$$\mathbf{q} = (z_1^2 - z_2^2, 2z_1 z_2). \tag{1.44}$$

We apply this to (1.40), satisfying $z_1(0) = \alpha > 0$, corresponding to the minor axis of an ellipse, where $w_1(0) = z_1'(0) = 0; z_2(0) = 0, w_2(0) = z_2'(0) = \lambda\beta$ (see Figure 1.6).

It is verified that

$$z_1(E) = \alpha \cos \frac{E}{2}, \quad z_2(E) = \beta \sin \frac{E}{2}, \tag{1.45}$$

where $E = 2\lambda s$ is the eccentric anomaly introduced in section 1.3. It is geometrically defined in Figure 1.6. The transformation of the ellipse $\mathbf{z}(E)$ by (1.44) into q_1, q_2 coordinates is shown in Figure 1.4. Equations (1.44), (1.45) imply

$$q_1(E) = -\frac{\beta^2 - \alpha^2}{2} + \frac{\beta^2 + \alpha^2}{2} \cos E, \quad q_2(E) = \alpha\beta \sin E.$$

These imply that the center of the ellipse is located at $q_1 = -\frac{\beta^2 - \alpha^2}{2}, q_2 = 0$. The semimajor axis $a = \frac{\beta^2 + \alpha^2}{2}$. Since the q_1-coordinate of the center of the ellipse, in absolute value, equals ae, then

$$ae = \frac{\beta^2 - \alpha^2}{2}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

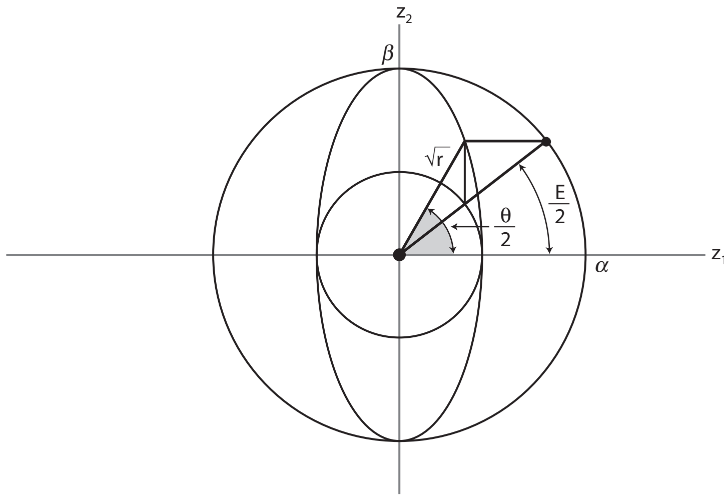

Figure 1.6 Geometry of eccentric anomaly.

and

$$\alpha = \sqrt{a(1-e)}, \beta = \sqrt{a(1+e)}.$$

Then

$$q_1(E) = a(\cos E - e), \quad q_2 = a\sqrt{1-e^2}\sin E. \tag{1.46}$$

Also

$$r(E) = a(1 - e\cos E), \tag{1.47}$$

$\dot{q}_1(E) = -\sqrt{ka}\ r^{-1}\sin E, \dot{q}_2(E) = \sqrt{ka}\ r^{-1}\sqrt{1-e^2}\cos E$. $t(E)$ is obtained from

$$t = \int r(s)ds = \frac{a}{2\lambda}\int (1 - e\cos E)dE$$

or

$$t = \frac{a}{2\lambda}(E - e\sin E), \tag{1.48}$$

where $\lambda = \sqrt{\frac{|h|}{2}} = \frac{1}{2}\sqrt{\frac{k}{a}}$. This yields $\mathcal{M} = E - e\sin E$. (1.48) is called *Kepler's equation*, and $\mathcal{M} = \frac{2\lambda}{a}t = k^{\frac{1}{2}}a^{-\frac{3}{2}}t$, as previously defined at the end of section 1.3.

The Levi-Civita transformation is two-dimensional in position space. Its generalization to three dimensions was initially developed by P. Kustaanheimo [130]. Referring to $q = z^2, q \in \mathbb{C}, z \in \mathbb{C}$ as the Levi-Civita transformation, it maps $\mathbf{R^2} \to \mathbf{R^2}$. The generalization by Kustaanheimo and Stiefel [131], called the *KS transformation*, maps $\mathbf{R^4} \to \mathbf{R^3}$. It is briefly described here (See [214]).

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

The KS transformation is given by

$$\mathbf{q} = \Lambda(\mathbf{z})\mathbf{z}, \tag{1.49}$$

$$\Lambda(\mathbf{z}) = \begin{pmatrix} z_1 & -z_2 & -z_3 & z_4 \\ z_2 & z_1 & -z_4 & -z_3 \\ z_3 & z_4 & z_1 & z_2 \end{pmatrix},$$

$\mathbf{z} = (z_1, z_2, z_3, z_4) \in \mathbf{R^4}, \mathbf{q} = (q_1, q_2, q_3) \in \mathbf{R^3}$. As with the Levi-Civita transformation, $|\mathbf{q}| = |\mathbf{z}|^2$, and $\mathbf{z} = \mathbf{0} \to \mathbf{q} = \mathbf{0}$, where collision between m_1, m_2 is for the three-dimensional Kepler problem defined by (1.20), with $\mathbf{q} \in \mathbf{R^3}, \mathbf{p} \in \mathbf{R^3}$.

The inverse map to (1.49) is multiple valued, and each point $\mathbf{q}$ is mapped onto a circle S^1 of radius $\sqrt{|\mathbf{q}|}$ on a plane through the origin of $\mathbf{R^4}$. This circle is called a *fiber* in $\mathbf{R^4}$, and fibers corresponding to different points in $\mathbf{R^3}$ do not intersect.

A canonical extension of (1.49) to momentum coordinates $\mathbf{p} \in \mathbf{R^3}$ is obtained as

$$\mathbf{p} = \frac{1}{|\mathbf{z}|^2}\Lambda(\mathbf{z})\mathbf{w}, \tag{1.50}$$

$\mathbf{w} \in \mathbf{R^4}$. Equations (1.49), (1.50) together with $dt = |\mathbf{z}|^2 ds$ are applied to (1.21), where H is again given by (1.19). An analogous system to (1.39) is obtained together with an induced bilinear form, which is an integral:

$$z_4 w_1 - z_3 w_2 + z_2 w_3 - z_1 w_4 = 0.$$

The KS transformation illustrates that an increase in dimension can cause difficulties in generalizing a two-dimensional regularization. This is the case for obtaining a global regularization. However, a local regularization to the three-dimensional case can present no dimension increase, such as Sundman's transformation considered in this section.

An n-dimensional regularization of the n-dimensional Kepler problem is described in section 1.6; it avoids the difficulties from which the KS transformation suffers of being multiple valued. This regularization is closely tied in with the basic Riemannian geometries of constant curvature.

1.5 THE RESTRICTED THREE-BODY PROBLEM: FORMULATIONS

Throughout most of this book, we will be interested in the three-body problem under special restrictions. This version of the three-body problem is derived in different coordinate systems and under various assumptions.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Consider the planar three-body problem in Jacobi coordinates (1.12), (1.13) for $\mathbf{q}(t), \mathbf{Q}(t)$. Recall that $\mathbf{q}(t)$ describes the motion of P_2 about P_1, and $\mathbf{Q}(t)$ describes the motion of P_3 relative to the center of mass of P_1, P_2, located at $\boldsymbol{\beta} = \nu^{-1}(m_1\mathbf{x}_1 + m_2\mathbf{x}_2)$, as shown in Figure 1.1 in an inertial coordinate system. The origin of this coordinate system is the center of mass $\boldsymbol{\rho}$ of the three mass particles.

We will regard P_1, P_2 as planetary objects, such as the Sun and Jupiter, respectively, or the Earth and Moon, which move in elliptical orbits. P_2 will be regarded as the smaller of P_1, P_2. In our two examples of P_1, P_2, the mass ratio $\mu = m_2/(m_1 + m_2)$ is significantly smaller than 1. For Jupiter and the Sun, $\mu = 0.001$, and for Moon and Earth, $\mu = 0.012$. If $P_1 =$ Earth and $P_2 =$ Sun, then $\mu = 0.000003$. Thus $\mu \ll 1$. More generally, we assume, unless otherwise notified, that $0 \le \mu < 1/2$. The assumption on P_3 is that it has negligible mass relative to P_1, P_2. For example, P_3 could be a small asteroid, comet, or spacecraft. Because P_1, P_2 are planetary-sized objects the gravitational perturbation on them due to P_3 will be negligible. Thus, P_1, P_2 can be viewed as a decoupled binary system, and their relative motion will reduce to the Kepler two-body problem (1.20). This is seen from (1.12) if we set $m_3 = 0$. By our assumptions, $m_3 \approx 0$. On the other hand, the motion of the particle P_3 of zero mass will be gravitationally perturbed by the two-body elliptical motion of P_1, P_2. The motion of P_3 is defined by (1.13), which yields a well-defined system as $m_3 \to 0$.

Setting $m_3 = 0$ implies $M = \nu$, and also

$$\boldsymbol{\beta} = \boldsymbol{\rho} = 0. \tag{1.51}$$

Thus, the origin of the coordinate system corresponds to the center of mass of P_1, P_2. Another restriction we make is to perform a time scaling so that $G = 1$. This scaling is given by $t \to \sqrt{G}t = \tilde{t}$, i.e., $d/dt = \sqrt{G}d/d\tilde{t}$. Thus, with $G = 1$, $\tilde{t}$ is the new time variable, and we relabel t for notational convenience. We also scale the masses m_1, m_2 so that $\nu = m_1 + m_2 = 1$. The mass is therefore dimensionless. With this normalization we set $m_1 = 1 - \mu, m_2 = \mu$, where $\mu = m_2/(m_1 + m_2)$.

As a final restriction, we assume that P_1, P_2 describe circular orbits, i.e., $e = 0$ in (1.28), implying $r = a = k/2|h|$ by (1.29), where $k = G\nu = 1$. Choosing the unit of length so that $r = |\mathbf{q}| = a = 1$ (i.e., $H = h = -\frac{1}{2}$) defines the *planar circular restricted three-body problem*, or *restricted problem* for short. Thus, it is defined in inertial coordinates by

$$\ddot{\mathbf{Q}} = -\frac{1-\mu}{r_{13}^3}(\mathbf{Q} - m_2\mathbf{q}(t)) + \frac{\mu}{r_{23}^3}(\mathbf{Q} - m_1\mathbf{q}(t)), \tag{1.52}$$

where $\mathbf{q}(t)$ defines the circular motion of P_2 about P_1 in an inertial coordinate system centered at P_1. The center of mass $\boldsymbol{\beta}$ of P_1, P_2 is at the origin, and according to (1.46),

$$q_1(t) = \cos t, \quad q_2(t) = \sin t. \tag{1.53}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Equation (1.48) implies $t = E = s$. The angular velocity of P_2 about P_1 is 1, and the period of motion is 2π. The relative distances of P_3 to P_1 and P_2 are

$$\begin{aligned} r_{13}(t) &= \sqrt{(Q_1 + \mu C)^2 + (Q_2 + \mu S)^2}, \\ r_{23}(t) &= \sqrt{(Q_1 - (1-\mu)C)^2 + (Q_2 - (1-\mu)S)^2} \end{aligned} \tag{1.54}$$

where $C \equiv \cos t, S \equiv \sin t$.

Definition 1.17 *Equation (1.52) defines the restricted problem in* inertial coordinates *whose origin is the center of mass of* P_1, P_2, *where* $m_1 = 1 - \mu, m_2 = \mu, 0 \le \mu < 1/2, r = |\mathbf{q}| = 1$.

Equation (1.52) can be written as

$$\ddot{\mathbf{Q}} = \Omega_{\mathbf{Q}}, \tag{1.55}$$

where

$$\Omega = \frac{1-\mu}{r_{13}} + \frac{\mu}{r_{23}}. \tag{1.56}$$

Ω is the potential energy of m_3, $\Omega = \Omega(\mathbf{Q}, t)$. For $\mu = 0$, the mass of P_2 vanishes, and (1.55) reduces to the two-body problem in inertial coordinates between P_3 and P_1 of mass $m_1 = 1$ at the origin. In this case (1.53) reduces to (1.20) with $G = \nu = 1$, with $\mathbf{q}$ replaced by $\mathbf{Q}$, and $r_{13} = |\mathbf{Q}|$.

Lemma 1.18 *The total energy* $\mathcal{H}$ *of the full system for* P_1, P_2, P_3 *for the restricted problem is*

$$\mathcal{H} = -\frac{m_1 m_2}{2}. \tag{1.57}$$

Proof. For the planar three-body problem, $\mathcal{H}$ is given by (1.16). Following the sequence of normalizations made for the restricted problem we set $m_3 = 0$, which implies $k_1 = 0, \mathbf{P} = \mathbf{0}, k_2 = m_1 m_2 \nu^{-1}$. With the normalization $G = 1, m_1 + m_2 = 1$, (1.16) reduces to

$$\mathcal{H} = \frac{1}{2} m_1 m_2 |\dot{\mathbf{q}}|^2 - \frac{m_1 m_2}{|\mathbf{q}|}.$$

Therefore by (1.29), where $H = h = -1/2$,

$$\mathcal{H} = m_1 m_2 \left[\frac{1}{2} |\dot{\mathbf{q}}|^2 - \frac{1}{|\mathbf{q}|} \right] = m_1 m_2 h = -\frac{m_1 m_2}{2}. \qquad \square$$

The restricted problem can be described as determining the motion of a particle of zero mass in a gravitational field generated by the uniform circular

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

motion of the mass points P_1, P_2, called the *primaries.* Since $m_3 = 0$, (1.57) can be viewed as the energy of P_1, P_2. It is an integral for the two-body motion of P_1, P_2.

The restricted problem possesses an integral associated to the energy of only P_3.

Definition 1.19 *The restricted problem (1.55) has an energy integral called the* Jacobi integral, *and in* inertial coordinates *it is given by*

$$J = J(\mathbf{Q}, \dot{\mathbf{Q}}, t) = -|\dot{\mathbf{Q}}|^2 + 2(Q_1\dot{Q}_2 - Q_2\dot{Q}_1) + 2\Omega. \tag{1.58}$$

The term $c = Q_1\dot{Q}_2 - Q_2\dot{Q}_1$ is the *angular momentum of* m_3, which in polar coordinates r, θ is equivalent to $c = r^2\dot{\theta}, r = \sqrt{Q_1^2 + Q_2^2}, \theta = \arctan(\frac{Q_2}{Q_1})$. J is alternatively referred to as the *Jacobi energy.*

Lemma 1.20 *J is an integral of (1.55).*

Proof. Let $\mathbf{Q} = \mathbf{Q}(t)$ represent a solution of (1.55). We need to show

$$\frac{d}{dt}J(\mathbf{Q}, \dot{\mathbf{Q}}, t) = 0.$$

Now,

$$\begin{aligned}\frac{d}{dt}J &= (J_{\mathbf{Q}}, \dot{\mathbf{Q}}) + (J_{\dot{\mathbf{Q}}}, \ddot{\mathbf{Q}}) + J_t \\ &= 2(\Omega_{\mathbf{Q}}, \dot{\mathbf{Q}}) - 2(\dot{\mathbf{Q}}, \ddot{\mathbf{Q}}) + 2(Q_1\ddot{Q}_2 - Q_2\ddot{Q}_1) + 2\Omega_t.\end{aligned}$$

By (1.55) the first two terms cancel, and

$$\frac{1}{2}\frac{d}{dt}J = Q_1\Omega_{Q_2} - Q_2\Omega_{Q_1} + \Omega_t.$$

We now verify that $Q_1\Omega_{Q_2} - Q_2\Omega_{Q_1} = -\Omega_t$, thus yielding the proof.

Using (1.56), direct calculation yields

$$\begin{aligned}&Q_1\Omega_{Q_2} - Q_2\Omega_{Q_1} \\ &= -\frac{\mu(1-\mu)}{r_{13}^3}(Q_1S - Q_2C) - \frac{\mu(1-\mu)}{r_{23}^3}(-Q_1S + Q_2C)\end{aligned}$$

and

$$\begin{aligned}\Omega_t &= -\frac{1-\mu}{r_{13}^3}\left[(Q_1 + \mu C)(-\mu S) + (Q_2 + \mu S)\mu C\right] \\ &\quad -\frac{\mu}{r_{23}^3}\left[(Q_1 - (1-\mu)C)(1-\mu)S - (Q_2 - (1-\mu)S)(1-\mu)C\right] \\ &= -\left[Q_1\Omega_{Q_2} - Q_2\Omega_{Q_1}\right].\end{aligned}$$

Therefore, $\dot{J} = 0$. □

Lemma 1.20 implies the following lemma.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Lemma 1.21 *The set*

$$J^{-1}(C) = \left\{(\mathbf{Q}, \dot{\mathbf{Q}}, t) \in \mathbf{R^5} | J = C, C \in \mathbf{R^1}\right\}$$

is a four-dimensional surface in the five-dimensional extended phase space on which the solution $\mathbf{Q}(t), \dot{\mathbf{Q}}(t)$ *lies for a given value of* C. *(See Figure 1.7.) (The constant* C *should not be confused with* $C = \cos t$*)*

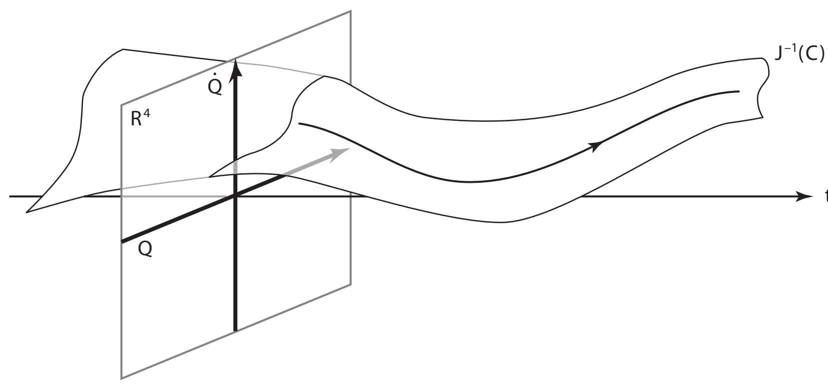

Figure 1.7 Jacobi integral surface.

Definition 1.22 C *is called the* Jacobi constant*, or* Jacobi energy*.*

Equation (1.55) can be transformed to a new coordinate system which is *time independent*. That is, it is an *autonomous system*. Since P_1, P_2 move about the origin with constant angular velocity of 1, then transforming to a new coordinate system x_1, x_2, which also rotates with an angular velocity ω of 1, implies that P_1, P_2 will be fixed. (See Figure 1.8.)

Without loss of generality, we assume that P_1, P_2 lie fixed on the x_1-axis.

Definition 1.23 *The* x_1, x_2*-coordinate system is called a* rotating coordinate system *or* fixed coordinate system.

The transformation $\mathbf{Q} \to \mathbf{x}$ is thus a uniform rotation given by an orthogonal matrix $\mathbf{R}(t)$,

$$\mathbf{x} = \mathbf{R}(t)\mathbf{Q}, \tag{1.59}$$

$$\mathbf{R}(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

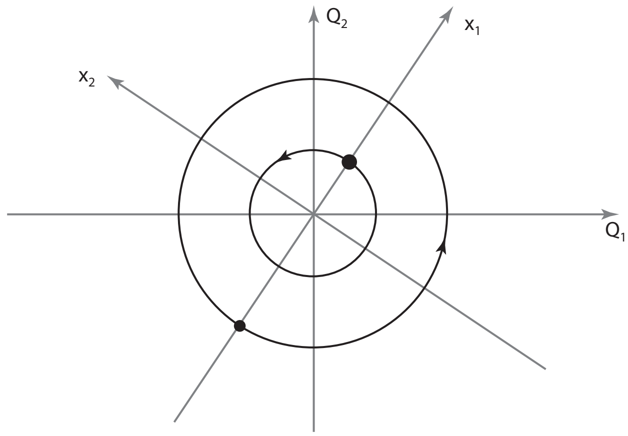

Figure 1.8 Rotating and inertial coordinates.

The inverse transformation is

$$\mathbf{Q} = \mathbf{R^{-1}}(t)\mathbf{x}, \tag{1.60}$$

where

$$R^{-1}(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

It is verified that substitution of (1.60) (i.e., $Q_1 = x_1 \cos t - x_2 \sin t, Q_2 = x_1 \sin t + x_2 \cos t$), into (1.55) yields the restricted problem in *rotating or fixed* coordinates,

$$\begin{aligned} \ddot{x}_1 - 2\dot{x}_2 &= x_1 + \tilde{\Omega}_{x_1}, \\ \ddot{x}_2 + 2\dot{x}_1 &= x_2 + \tilde{\Omega}_{x_2}, \end{aligned} \tag{1.61}$$

where

$$\tilde{\Omega} = \tilde{\Omega}(\tilde{x}) = \Omega(\mathbf{R^{-1}x}) = \frac{1-\mu}{r_{13}} + \frac{\mu}{r_{23}}, \tag{1.62}$$

$$r_{13} = \sqrt{(x_1 - \mu)^2 + x_2^2}, \quad r_{23} = \sqrt{(x_1 - (-1+\mu))^2 + x_2^2}.$$

The particle P_1 is fixed at $\mathbf{x} = (\mu, 0)$, and P_2 is fixed at $\mathbf{x_2} = (-1 + \mu, 0)$. (See Figure 1.9.) The system (1.61) is a standard form of the restricted problem commonly seen in the literature.

The system (1.61) is an autonomous system of differential equations.

When $\mu = 0$, (1.61) reduces to the two-body problem in rotating coordinates between P_3 of zero mass and P_1 at the origin of unit mass, and $r_{13} = |\mathbf{x}| = |\mathbf{Q}|$. The system (1.62) has the same form as (1.56), except t is not present, and r_{13}, r_{23} are simplified.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

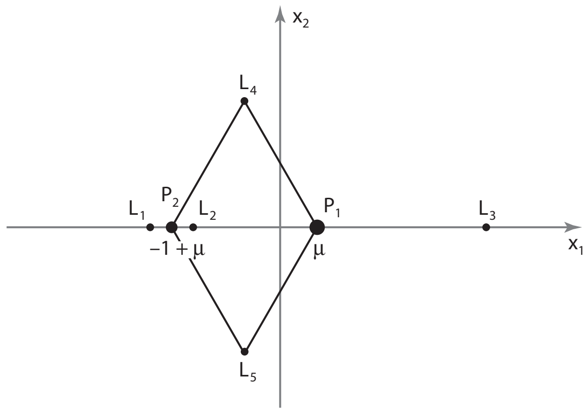

Figure 1.9 Rotating coordinate system showing Lagrange points.

The right-hand side of (1.61) represents the sum of two forces $\mathbf{F}, \mathbf{G}$,

$$\mathbf{F} = (x_1, x_2), \qquad \mathbf{G} = (\tilde{\Omega}_{x_1}, \tilde{\Omega}_{x_2}).$$

$\mathbf{F}$ is the outward radially directed centrifugal force, and $\mathbf{G}$ is the sum of the gravitational forces due to m_1, m_2 on m_3. $\mathbf{F}$ exists due to the fact of being in a rotating coordinate system and is an artificially induced force.

By direct substitution of (1.60) into (1.58) we obtain the Jacobi integral in rotating coordinates. It is verified that

$$\tilde{J}(\mathbf{x}, \dot{\mathbf{x}}) = J(\mathbf{R^{-1}}(t)\mathbf{x}, \quad \frac{d}{dt}\mathbf{R^{-1}}(t)\mathbf{x}) = -|\dot{\mathbf{x}}|^2 + r^2 + 2\tilde{\Omega}(\mathbf{x}), \qquad (1.63)$$

where $r = |\mathbf{x}| = |\mathbf{R}(t)\mathbf{Q}| = |\mathbf{Q}|$. It is verified by direct differentiation of (1.63) using (1.61) that indeed $\frac{d}{dt}\tilde{J} = 0$.

Definition 1.24 $\tilde{J}(\mathbf{x}, \dot{\mathbf{x}})$ *is called the Jacobi integral in* rotating coordinates.

Lemma 1.25 *The set*

$$\tilde{J}^{-1}(C) = \{(\mathbf{x}, \dot{\mathbf{x}}) \in \mathbf{R^4} | \tilde{J} = C\} \qquad (1.64)$$

is a three-dimensional surface in the four-dimensional phase space on which the solutions $x(t), \dot{x}(t)$ of (1.61) lie for a given value of C.

Thus, the dimension of the Jacobi integral manifold for the given value of C has decreased by 1 when changing from the Q_1, Q_2 coordinates to the x_1, x_2 coordinates.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

The Jacobi integral (1.58) for (1.55) can be written in a more standardized form,

$$\hat{J} = \frac{1}{2}|\dot{\mathbf{Q}}|^2 - (Q_1\dot{Q}_2 - Q_2\dot{Q}_1) - \Omega(\mathbf{Q}, t) \tag{1.65}$$

or

$$\hat{J} = E - c, \tag{1.66}$$

where

$$E = \frac{1}{2}|\dot{\mathbf{Q}}|^2 - \Omega(\mathbf{Q}, t) \tag{1.67}$$

is the total energy of P_3 in the barycentric inertial system (i.e., where the center of mass is at the origin) and

$$c = Q_1\dot{Q}_2 - Q_2\dot{Q}_1 \tag{1.68}$$

is the angular momentum. For $\mu \neq 0$ neither E nor c are integrals. They are both integrals when $\mu = 0$, where $m_2 = 0$ and E reduces to the two-body problem between P_1 of mass $m_1 = 1$ and P_3. In that case, E reduces to (1.19).

It is verified that

$$L(\mathbf{Q}, \dot{\mathbf{Q}}) = L(\mathbf{x}, \dot{\mathbf{x}}) + |\mathbf{x}|^2, \tag{1.69}$$

where $L(\mathbf{x}, \dot{\mathbf{x}}) = \dot{x}_1 x_2 - \dot{x}_2 x_1$. $L(\mathbf{Q}, \dot{\mathbf{Q}})$ is an integral for the two-body problem (1.18) with $\mathbf{q}$ replaced by $\mathbf{Q}$. This problem has energy (1.19), which we more generally write as

$$H = \frac{1}{2}|\dot{\mathbf{Q}}|^2 - \frac{a}{|\mathbf{Q}|},$$

where $a \in \mathbf{R^1}$ is a constant. Since $L(\mathbf{Q}, \dot{\mathbf{Q}})$ is an integral for the Kepler flow defined by (1.18) in inertial coordinates, then $L(\mathbf{x}, \dot{\mathbf{x}}) + |\mathbf{x}|^2$ is an integral for the Kepler flow in rotating coordinates, given by (1.61) with $\mu = 0$, where $r_{13} = |\mathbf{x}|$.

It is also verified by (1.60) that (1.19), with $G\nu$ replaced by a, in rotating coordinates is

$$\tilde{H}(\mathbf{x}, \dot{\mathbf{x}}) = H(R^{-1}\mathbf{x}, \frac{d}{dt}(R^{-1}\mathbf{x})) = \frac{1}{2}|\dot{\mathbf{x}}|^2 + \frac{1}{2}|\mathbf{x}|^2 - L(\mathbf{x}, \dot{\mathbf{x}}) - \frac{a}{|\mathbf{x}|}, \tag{1.70}$$

which is therefore an integral of the Kepler flow in rotating coordinates. Thus, in summary, we have the following lemma.

Lemma 1.26

$$\tilde{L} = L(\mathbf{x}, \dot{\mathbf{x}}) + |\mathbf{x}|^2,$$
$$\tilde{H} = \frac{1}{2}|\dot{\mathbf{x}}|^2 + \frac{1}{2}|\mathbf{x}|^2 - L(\dot{\mathbf{x}}, \dot{\mathbf{x}}) - \frac{a}{|\mathbf{x}|},$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

corresponding to the angular momentum and energy of the two-body problem in rotating coordinates, are integrals of the Kepler flow in rotating coordinates,

$$\ddot{x}_1 - 2\dot{x}_2 = x_1 - \frac{a}{|\mathbf{x}|^3},$$
$$\ddot{x}_2 - 2\dot{x}_1 = x_2 - \frac{a}{|\mathbf{x}|^3}.$$

It is noted that

$$\tilde{J}|_{\mu=0} = -2\left[\tilde{H} - \tilde{L}\right], \tag{1.71}$$

where $a = 1$. More generally we calculate $\tilde{J}$ for $\mu \neq 0$ in terms of $\tilde{H}$ in P_2-centered coordinates in chapter 3, section 3.2.

Equation (1.55) can be written in Hamiltonian form:

$$\dot{\mathbf{Q}} = E_{\mathbf{P}},\ \dot{\mathbf{P}} = -E_{\mathbf{Q}}, \tag{1.72}$$

where E is given by (1.67) with $\dot{\mathbf{Q}} = \mathbf{P}$. E is not an integral for $\mu \neq 0$ because it is time dependent: $\dot{E} = (\mathbf{P}, \dot{\mathbf{P}}) - (\Omega_{\mathbf{Q}}, \dot{\mathbf{Q}}) - \Omega_t = -\Omega_t \neq 0$.

Likewise, for rotating coordinates, (1.61) can be written as a Hamiltonian system,

$$\dot{\mathbf{q}} = G_{\mathbf{p}}, \qquad \dot{\mathbf{p}} = -G_{\mathbf{q}}, \tag{1.73}$$

where

$$G = \frac{1}{2}|\mathbf{p}|^2 - (q_1 p_2 - q_2 p_1) - \tilde{\Omega}(\mathbf{q}) \tag{1.74}$$

and G is an integral for (1.73) since it is time independent.

Lemma 1.27 *Equation (1.61) is obtained from (1.73) by the map*

$$q_k = x_k, \qquad p_1 = \dot{x}_1 - x_2, \qquad p_2 = \dot{x}_2 + x_1, \tag{1.75}$$

and

$$G(\mathbf{p}, \mathbf{q}) = -\frac{1}{2}\tilde{J}(\mathbf{x}, \dot{\mathbf{x}}).$$

Proof.

$$\dot{\mathbf{q}} = \mathbf{p} + \begin{pmatrix} q_2 \\ -q_1 \end{pmatrix},$$
$$\dot{\mathbf{p}} = \begin{pmatrix} p_2 \\ -p_1 \end{pmatrix} - \tilde{\Omega}_{\mathbf{q}}(\mathbf{q}).$$

Therefore, these yield, respectively,

$$\dot{\mathbf{x}} = \begin{pmatrix} \dot{x}_1 - x_2 \\ \dot{x}_2 + x_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

and

$$\begin{pmatrix} \ddot{x}_1 - \dot{x}_2 \\ \ddot{x}_2 + \dot{x}_1 \end{pmatrix} = \begin{pmatrix} \dot{x}_2 + x_1 \\ -\dot{x}_1 + x_2 \end{pmatrix} - \tilde{\Omega}_{\mathbf{x}}(\mathbf{x}),$$

which yields (1.61).

$$\begin{aligned} &G(\mathbf{p}(\mathbf{x}, \dot{\mathbf{x}}), \mathbf{q}(\mathbf{x}, \dot{\mathbf{x}})) = \tilde{G}(\mathbf{x}, \dot{\mathbf{x}}) \\ &= \frac{1}{2}[(\dot{x}_1 - x_2)^2 + (\dot{x}_2 + x_1)^2] - [x_1(\dot{x}_2 + x_1) - x_2(\dot{x}_1 - x_2)] - \tilde{\Omega}(\mathbf{x}) \\ &= \frac{1}{2}|\dot{\mathbf{x}}|^2 - \dot{x}_1 x_2 + \dot{x}_2 x_1 + \frac{1}{2}|\mathbf{x}|^2 - x_1\dot{x}_2 + x_2\dot{x}_1 - |x|^2 - \tilde{\Omega}(\mathbf{x}). \end{aligned}$$

Thus,

$$\tilde{G} = \frac{1}{2}|\dot{\mathbf{x}}|^2 - \frac{1}{2}r^2 - \tilde{\Omega}(\mathbf{x}) = -\frac{1}{2}\tilde{J}. \tag{1.76}$$

□

Note that although G is a Hamiltonian function for (1.73), neither $\tilde{G}$ nor $\tilde{J}$ are Hamiltonian functions for (1.61), but they are integrals for that system. Equation (1.61) is not in canonical form since (1.75) is not a canonical map. This follows since

$$\begin{aligned} (\mathbf{dp}, \mathbf{dq}) &= (dy_1 - dx_2)dx_1 + (dy_2 + dx_1)dx_2 \\ &= (\mathbf{dy}, \mathbf{dx}) + 2dx_1 dx_2, \end{aligned}$$

where we used the antisymmetry of the products of one-forms dx, dy $dxdy = -dydx$ (see [15]). Thus $(\mathbf{dp}, \mathbf{dq}) \neq (\mathbf{dy}, \mathbf{dx})$.

Definition 1.28 *The* three-dimensional restricted problem *in* inertial coordinates *is given by (1.55), (1.56) with* $\mathbf{Q} = (Q_1, Q_2, Q_3) \in \mathbf{R^3}$*, and*

$$\begin{aligned} r_{13} &= \sqrt{(Q_1 + \mu\cos t)^2 + (Q_2 + \mu \sin t)^2 + Q_3^2}, \\ r_{23} &= \sqrt{(Q_1 - (1-\mu)\cos t)^2 + (Q_2 - (1-\mu)\sin t)^2 + Q_3^2}, \end{aligned}$$

where J is given by (1.58). The transformation to rotating coordinates $\mathbf{x} = (x_1, x_2, x_3) \in \mathbf{R^3}$ *is given by*

$$\mathbf{x} = S(t)\mathbf{Q},$$

where

$$S(t) = \begin{pmatrix} \mathbf{R}(t) & 0 \\ 0 & 1 \end{pmatrix},$$

leaving the Q_3-axis invariant. This maps (1.55) into the system given by (1.61) together with $\ddot{x}_3 = \tilde{\Omega}_{x_3}$, where $r_{13} = \sqrt{(x_1 - \mu)^2 + x_2^2 + x_3^2}, r_{23} = \sqrt{(x_1 - (-1+\mu))^2 + x_2^2 + x_3^2}$. The Jacobi integral is given by (1.63), where $r = |\mathbf{x}| = |\mathbf{S}(t)\mathbf{Q}|$.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

A final definition for this section is for the *three-dimensional elliptic restricted three-body problem* in inertial coordinates. It parallels the derivation of (1.52), (1.55). We begin with (1.11) with $\mathbf{x}_k \in \mathbf{R^3}$, and obtain (1.12), (1.13), where $\mathbf{q} \in \mathbf{R^3}, \mathbf{Q} \in \mathbf{R^3}$. This represents the three-dimensional three-body problem in Jacobi coordinates. Setting $m_3 = 0$, the center of mass of P_1, P_2 is at the origin, expressed in (1.51); the time is scaled so that $G = 1$; and $m_1 + m_2 = 1, m_1 = 1 - \mu, m_2 = \mu, 0 \leq \mu < 1/2$. Instead of choosing $e = 0$ in (1.28), it is more generally assumed that $e \in [0, 1)$, where r is given by (1.28) for $h < 0$, expressed as $r = r(\theta), \theta \in [0, 2\pi]$. The unit of length is normalized so that the semimajor axis $a = 1$, or equivalently $h = -\frac{1}{2}$.

For the binary system m_1, m_2, since the motion is planar it is assumed without loss of generality to lie in the q_1, q_2-plane so that for the elliptic motion of this pair, $q_3 = 0$. Thus, in (1.12) $\mathbf{q} = \mathbf{q}(t) = (q_1(t), q_2(t), 0)$, where $q_1(t), q_2(t)$ are defined by (1.18), with $G\nu = 1$.

Equation (1.55) is again obtained with two changes. First, $\mathbf{Q} \in \mathbf{R^3}$ and r_{13}, r_{23} take a different form.

Definition 1.29 *The three-dimensional elliptic restricted three-body problem is defined by*

$$\ddot{\mathbf{Q}} = \Omega_{\mathbf{Q}}, \tag{1.77}$$

$\mathbf{Q} \in \mathbf{R^3}$, *with*, $\mathbf{\Omega}$ *given by (1.56), where*

$$r_{13} = |\mathbf{Q} + \mu\mathbf{q}(t)|, r_{23} = |\mathbf{Q} - (1 - \mu)\mathbf{q}(t)|; \tag{1.78}$$

$\mathbf{q}(t) = (q_1(t), q_2(t), 0), q_k(t), k = 1, 2$, *are given by (1.46) with* $a = 1$ *and* E *is given implicitly by (1.48).*

Equation (1.77) reduces to the three-dimensional restricted problem when $e = 0$. Thus, the three-dimensional elliptic restricted problem in Definition 1.29 describes the motion of a zero mass particle P_3 in a Newtonian gravitational field generated by the Keplerian elliptic motion of the binary pair P_1, P_2.

If $0 < \mu \ll 1$ P_1 lies very close to the origin and P_2 moves about the center of mass of P_1, P_2, which lies very close to P_1, then P_2 moves approximately about P_1, approximately at the origin. If, moreover, $e \gtrsim 0$, then the situation approximates the motions of most of the planets of our solar system moving about the Sun. Notable exceptions are Pluto and Mercury, where $e = 0.247, 0.206$, respectively. Also, evidence is pointing to the fact that Pluto is not a planet but rather belongs to a different class of objects called Kuiper belt objects [41, 21]. For all the other planets $e \doteq 0.0n, n \in \{0, 1, \ldots, 9\}$. Also note that the planetary ephemeris which accurately describes the motions of the planets [212], e, as well as the other

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

orbital parameters, are not actually constant, and have small variations. This is due primarily to the fact that the Sun and planets are not mass points and are spherical in shape, with finite radius. Also, they are not purely spherical and have oblateness perturbations which distort the ideal gravitational field. Moreover, all the planets gravitationally interact with one another, and their orbits are gravitationally perturbed by neighboring stars. In addition, our solar system itself orbits the center mass of the Milky Way galaxy. There are also non-gravitational perturbations, and forces not understood or known.

Nevertheless, the variations in the orbital parameters of the planets are negligible and the ideal model we are using by viewing the Sun and planets as point masses yields accurate results when compared to reality in many situations.

The existence of solutions to the planar restricted problem in rotating coordinates (1.61) is addressed in chapter 2, where quasi-periodic motion is proven to exist by the Kolmogorov-Arnold-Moser(KAM) theorem. Reduction of the flow of the restricted problem to a monotone twist map proves the existence of chaotic motion due to the complicated intersection of hyperbolic invariant manifolds, and also by applying Aubrey-Mather theory.

In chapter 3, different types of capture solutions for the restricted problem are proven to exist and parabolic orbits are studied. Some of the capture solutions lie near parabolic orbits and are proven to be associated to a hyperbolic invariant set which gives rise to chaotic motion. Applications are also discussed.

1.6 THE KEPLER PROBLEM AND EQUIVALENT GEODESIC FLOWS

In this section we prove that the flow of the Kepler problem is equivalent to the geodesic flow of the basic Riemannian spaces of constant Gaussian curvature K. These spaces turn out to be topologically equivalent to the fixed energy surfaces of the Kepler problem. This is done for $n \geq 2$ dimensions and provides an n-dimensional regularization of collision in the Kepler problem. The regularized differential equations are just the differential equations for the geodesic flows and are easily solved explicitly. The proof of this result follows from Moser [174], Osipov [182, 183], and Belbruno [22, 24].

We will describe the basic results for the three Kepler energies, $H < 0, H > 0, H = 0$, where H is the Kepler energy. These results will be

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

described separately for each respective energy case, and the basic theorems stated. We will then give a new proof for all the cases at once. The n-dimensional Kepler problem is defined by

$$\ddot{\mathbf{q}} = -\frac{\mathbf{q}}{|\mathbf{q}|^3}, \tag{1.79}$$

$\cdot \equiv \frac{d}{dt}, \mathbf{q} = (q_1, q_2, \ldots, q_n) \in \mathbf{R^n}$, which can be written as a Hamiltonian system

$$\dot{\mathbf{q}} = H_{\mathbf{p}}, \quad \dot{\mathbf{p}} = -H_{\mathbf{q}}, \tag{1.80}$$

$\mathbf{p} = (p_1, p_2, \ldots, p_n) \in \mathbf{R^n}, n \geq 2,$

$$H = \frac{1}{2}|\mathbf{p}|^2 - \frac{1}{|\mathbf{q}|}. \tag{1.81}$$

Equations (1.79), (1.80) are just a generalization of (1.20), (1.21) to n dimensions. We define the $(2n-1)$-dimensional energy surface

$$H^{-1}(h) = \left\{(\mathbf{p}, \mathbf{q}) \in \mathbf{R^{2n}} | H = h, h \in \mathbf{R^1}\right\}. \tag{1.82}$$

We consider the three basic cases, $h = -\frac{1}{2}, \frac{1}{2}, 0$. The case of general $h < 0, h > 0$ follows by a simple scaling of $h = \frac{1}{2}, \frac{1}{2}$, respectively.

1.6.1 Case $h = -\frac{1}{2}$

In this case, the Kepler problem turns out to be topologically equivalent with the geodesic flow on the unit sphere, $S^n = \left\{\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_n) \in \mathbf{R^{n+1}} | |\boldsymbol{\xi}|^2 = \sum_{k=0}^n \xi_k^2 = 1\right\}$. The equivalence is accomplished by the stereographic projection

$$p_k = \frac{\xi_k}{1 - \xi_0}, \quad k = 1, 2, \ldots, n, \tag{1.83}$$

$\xi_0 \neq 1$. See Figure 1.10.

This maps the great circles on $S_0^n = S^n \backslash \{\xi_0 = 1\}$, which are the geodesics, onto circles in $\mathbf{p}$-space. $\{\xi_0 = 1\} \equiv \{\xi_0 = 1, \xi_1 = \xi_2 = \cdots = \xi_n = 0\}$ is the north pole of the sphere. The notation $S^n \backslash \{\xi_0 = 1\}$ means that the north pole point is deleted from the sphere. The circles in $\mathbf{p}$-space are called *hodographs* and correspond to the paths traced out by the velocity components of an elliptical path in $\mathbf{q}$-space. This type of equivalence first goes back to Fock [83], who applied it to the momentum variables to transform the Schrödinger wave equation in quantum mechanics. This was done for the case of S^3. The use of the projection (1.83) for Kepler's problem, also for $n = 3$, goes back to Györgyi [90], who canonically extended the map (1.83) to a mapping $q_k = q_k(\boldsymbol{\xi}, \boldsymbol{\eta}), \boldsymbol{\eta} \in \mathbf{R^4}$, where $\boldsymbol{\eta} = (\eta_0, \eta_1, \eta_2, \eta_3)$ is the velocity vector of a geodesic on S^3, therefore normal to $\boldsymbol{\xi}$. A general equivalence of the Kepler problem with the geodesic flow on $S^n, n \geq 2$, for $h < 0$ was done by Moser [174].

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

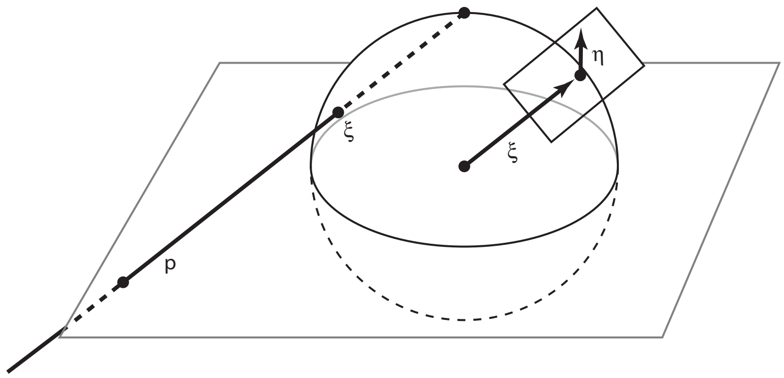

Figure 1.10 Stereographic projection from S^2.

The canonical extension of (1.83) in [174] is found by satisfying the relationship

$$\langle \boldsymbol{\eta}, \mathbf{d}\boldsymbol{\xi} \rangle = \langle \mathbf{y}, \mathbf{dx} \rangle, \tag{1.84}$$

$\mathbf{x} \in \mathbf{R^n}, \mathbf{y} \in \mathbf{R^n}, \boldsymbol{\eta} = (\eta_0, \eta_1, \ldots, \eta_n) \in \mathbf{R^{n+1}}$ and then setting $\mathbf{y} = \mathbf{q}, \mathbf{x} = -\mathbf{p}, \mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{R^n}, \mathbf{y} = (y_1, \ldots, y_n) \in \mathbf{R^n}$, $\langle \mathbf{y}, \mathbf{dx} \rangle \equiv \sum_{k=1}^n y_k dx_k$, $\langle \boldsymbol{\eta}, \mathbf{d}\boldsymbol{\xi} \rangle \equiv \sum_{k=0}^n \eta_k d\xi_k$. The variables $(\boldsymbol{\xi}, \boldsymbol{\eta}) \in T_1(S^n)$,

$$T_1(S^n) = \left\{ (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbf{R^{2n+2}} || \boldsymbol{\xi}| = 1, |\boldsymbol{\eta}| = 1, \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = 0 \right\}, \tag{1.85}$$

where $|\boldsymbol{\xi}|^2 = \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle$. $T_1(S^n)$ is called the *unit tangent bundle of* S^n. This means that all tangent vectors to S^n have unit length. *Tangent bundle* means the collection of all vectors in the tangent space of S^n.

It is found after simplification that

$$q_k = \eta_k(\xi_0 - 1) - \xi_k \eta_0, \tag{1.86}$$

$k = 1, 2, \ldots, n$. The geodesic flow on $T(S^n)$ is given by the Hamiltonian system

$$\boldsymbol{\xi}' = \Phi_{\boldsymbol{\eta}}, \quad \boldsymbol{\eta}' = -\Phi_{\boldsymbol{\xi}}, \tag{1.87}$$

where $' \equiv \frac{d}{ds}$,

$$\Phi = \frac{1}{2} |\boldsymbol{\xi}|^2 |\boldsymbol{\eta}|^2,$$

and Φ has the value of $\frac{1}{2}$. The time variable s is related to t by

$$t = \int |\mathbf{q}| ds, \tag{1.88}$$

which we also used in Section 1.4 for the Levi-Civita map, equation (1.36). System (1.87) yields the harmonic oscillator

$$\boldsymbol{\xi}'' + \boldsymbol{\xi} = \mathbf{0}. \tag{1.89}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

The solutions to this system, where $\eta = \xi'$, yield the geodesic curves $\xi(s)$ on S^n. It was proven in [174] that (1.83), (1.86), (1.88) on $T_1(S_0^n)$ transform the Kepler problem (1.80) into (1.87). The energy surface $H^{-1}(-\frac{1}{2})$ is then topologically equivalent to $T_1(S_0^n)$. The north pole $\{\xi_0 = 1\}$ corresponds to collision, as one sees from (1.83) since $\{\xi_0 = 1\} \to |\mathbf{p}| = \infty$. The geodesic on S_0^n passing through the north pole therefore corresponds to a collision orbit. Restoring the north pole to the punctured sphere S_0^n, where the transformed collision orbit is smooth, therefore regularizes the flow. It is noted that a geodesic passing through the north pole $\boldsymbol{\xi}^+$ must also pass through the south pole $\boldsymbol{\xi}^-$. The south pole projects onto the origin $\mathbf{p} = \mathbf{0}$.

The above equations are for the case $h = -\frac{1}{2}$, i.e., $H = -\frac{1}{2}$. The case of arbitrary negative energy $H = h = -\frac{a}{2}, a > 0$, is $\mathbf{q} \to a\mathbf{q}, \mathbf{p} \to a^{-\frac{1}{2}}\mathbf{p}, t \to a^{\frac{3}{2}}t$.

Theorem 1.30 (Moser) *The energy surface $H^{-1}(h), h < 0$, is topologically equivalent to $T_1(S_0^n)$, where the north pole corresponds to collision states. The geodesic flow on S_0^n is mapped into the Kepler flow after a change of the independent variable.*

The idea of the proof of Theorem 1.30 outlined above relies on Hamiltonian formalism.

1.6.2 Case $h = +\frac{1}{2}$

Following the general approach of Moser in the case of $h = -1/2$, the case $h = 1/2$ was first solved in [22] along with the case of $h = 0$ and for the central repelling field. A similar construction for these cases can be found in [183]. A survey article by Milnor [163] gives a geometric description of the solutions of the various cases.

The sphere has Gaussian curvature $K(S^n) = 1$ [180]. The Gaussian curvature is a real-valued function which measures the curvature of a manifold. A formula for this is given below.

Thus, it would seem reasonable that for positive energy a topologically equivalent surface for the geodesic flow for Kepler's problem would have a constant negative curvature, $K = -1$. This is indeed the case.

We summarize the results in [22] for the cases $h = 1/2, 0$ and for the central repelling field.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Instead of S^n, we have an n-dimensional hyperboloid of two sheets $H^n_{+,-}$ ($+,-$ refer to the upper and lower sheets, respectively).

$$H^n_{+,-} = \{\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_n) \in \mathbf{R^{n+1}} | \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle = \sum_{k=1}^{n} \xi_k^2 - \xi_0^2 = -1\},$$

embedded in a Lorentz space $\mathbf{L^{n+1}}$ defined by the metric

$$ds^2 = \sum_{k=1}^{n} d\xi_k^2 - d\xi_0^2, \tag{1.90}$$

$\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle \equiv \sum_{k=1}^{n} \xi_k \eta_k - \xi_0 \eta_0$. It is verified that $K(H^n_{+,-}) = -1$ and that ds^2 is Riemannian when restricted to $H^n_{+,-}$; i.e., it is positive definite.

Proceeding as in the negative energy case, we find that the geodesic flow is defined by the Hamiltonian system,

$$\boldsymbol{\xi}' = \Lambda \Phi_{\boldsymbol{\eta}}, \quad \boldsymbol{\eta}' = -\Lambda \Phi_{\boldsymbol{\xi}}, \tag{1.91}$$

where $\boldsymbol{\eta} = (\eta_0, \eta_1, \ldots, \eta_n)$,

$$\Lambda = \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & -1 \end{pmatrix},$$

which is the identity matrix except the lower right element is -1. It is verified that (1.91) is just

$$\boldsymbol{\xi}'' - \boldsymbol{\xi} = \mathbf{0}, \tag{1.92}$$

yielding hyperbolic solutions. These are the geodesics on $H^n_{+,-} \subset \mathbf{L^{n+1}}$ and are *great hyperbolas*, obtained by intersecting $H^n_{+,-}$ with any plane passing through the origin of $\boldsymbol{\xi}$-space.

We restrict ourselves to H^n_+ and restrict $(\boldsymbol{\xi}, \boldsymbol{\eta})$ to the unit tangent bundle of H^n_+,

$$\begin{aligned} T_1(H^n_+) = \big\{ & (\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbf{R^{2n+2}} | \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \\ & = -1, \langle \boldsymbol{\eta}, \boldsymbol{\eta} \rangle = 1, \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = 0 \big\}. \end{aligned}$$

The lower sheet H^n_- could have been used as well, but we use just one sheet to achieve a one-to-one correspondence with the Kepler problem. The map (1.83) is used to map $\boldsymbol{\xi} \in H^n_+$ to $\mathbf{p} \in \mathbf{R^n}$, where $\boldsymbol{\xi} \neq \boldsymbol{\xi}^+ = (1, 0, \ldots, 0)$, corresponding to the minimum point of H^n_+. Thus, $\boldsymbol{\xi} \in H^n_+ \backslash \boldsymbol{\xi}^+ \equiv H^n_{+,0}$. The map $\boldsymbol{\xi} \to \mathbf{p}$ is geometrically shown in Figure 1.11.

$\boldsymbol{\xi}^+$ is analogous to the north pole of S^n. It is seen that the great hyperbolas are mapped one-to-one onto a circle in $\mathbf{p}$-space, where part of the circle

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

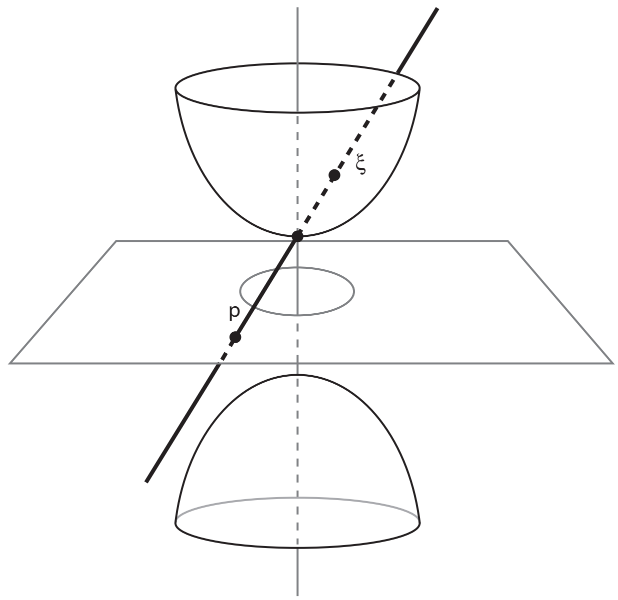

Figure 1.11 Projection from two-sheeted hyperboloid.

is missing. (See Figure 1.12.) This circle turns out to correspond to the velocity curve traced out for the Kepler problem for a hyperbolic trajectory in **q**-space. This is analogous to the case of elliptic motion. Hamilton [95] and Möbius [164] classified the velocity hodographs for the Kepler problem. They defined the term *velocity hodograph* to represent the velocity curve $\mathbf{p}(t)$ corresponding to a Kepler orbit in $\mathbf{q}(t)$.

Theorem 1.31 (Möbius, Hamilton) *The velocity hodographs for Kepler's problem are circles or parts of circles.*

It is seen that (1.83) maps $H^n_{+,0}$ isometrically into the space of $\hat{D}^n, |\mathbf{p}| > 1$, where isometric means that the induced metric ds^2 on $H^n_+ \subset \mathbf{L^{n+1}}$ is preserved in $\hat{D}^n$. $\hat{D}^n$ is a metric space which has the velocity hodographs as geodesics, which intersect $\partial\hat{D}^n, |\mathbf{p}| = 1$, normally. See Figure 1.12.

The metric on $\hat{D}^n$ is

$$ds^2 = 4(|\mathbf{p}|^2 - 1)^{-2}|\mathbf{dp}|^2, \tag{1.93}$$

which implies $K(\hat{D}^n) = 1$. This must be the case since (1.83) is isometric, and $K(H^n_+) = 1$.

A canonical extension of (1.83) is given by

$$q_k = \eta_k(\xi_0 - 1) - \xi_k\eta_0, \tag{1.94}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

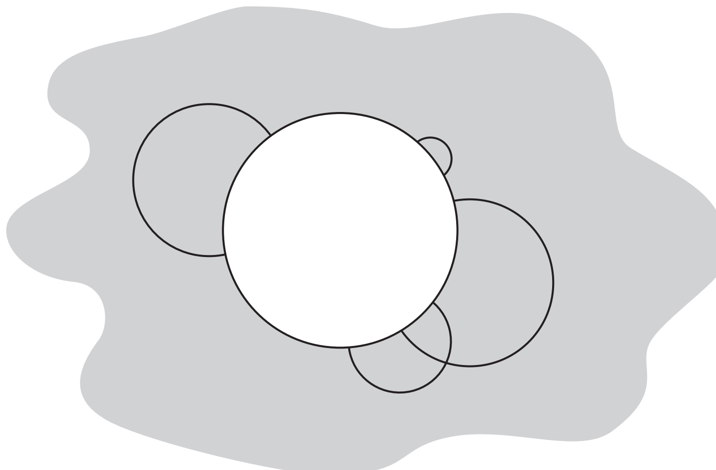

Figure 1.12 Inverted Lobachevsky disc and geodesics.

$k = 1, 2, \ldots, n$, which satisfies

$$\langle \boldsymbol{\eta}, \mathbf{d}\boldsymbol{\xi} \rangle = \langle \mathbf{y}, \mathbf{dx} \rangle, \tag{1.95}$$

$\mathbf{x} \in \mathbf{R^n}, \mathbf{y} \in \mathbf{R^n}$, and where $\mathbf{y} = \mathbf{q}, \mathbf{x} = -\mathbf{p}$. Equations (1.83), (1.94) together with (1.88) maps (1.91) into the Kepler problem (1.80) and the energy surface $H^{-1}(\frac{1}{2})$ is topologically equivalent to $T_1(H^n_{+,0})$. Thus, as stated in [22], we have the following theorem.

Theorem 1.32 *The energy surface $H^{-1}(h), h > 0$, is topologically equivalent to $T_1(H^n_{+,0}), H^n_+ \subset \mathbf{L^{n+1}}$, where the point ξ^+ corresponds to collision. The geodesic flow on $H^n_{+,0}$ is mapped one-to-one into the Kepler flow after a change of the independent variable.*

Regularization of collision is achieved in (1.91) by restoring ξ^+ to $H^n_{+,0}$ where (1.91) is smooth. Great hyperbolas passing through ξ^+ correspond to collision orbits. The case of general energy $H = h = \frac{a}{2} > 0$ is accomplished by the scaling $\mathbf{q} \to a\mathbf{q}, \mathbf{p} \to a^{-\frac{1}{2}}\mathbf{p}, t \to a^{\frac{3}{2}}t$.

It is interesting to note that inversion with respect to $\hat{D}^n, \tilde{p}_k = p_k/|\mathbf{p}|^2, k = 1, 2, \ldots, n, |\tilde{\mathbf{p}}| = |\mathbf{p}|^{-1}$, is an isometry, and when applied to (1.83) yields the map

$$\tilde{p}_k = \frac{\xi_k}{1 + \xi_0}, \qquad k = 1, 2, \ldots, n. \tag{1.96}$$

$\hat{D}^n$ is mapped into $D^n = \{|\tilde{\mathbf{p}}| < 1\}$ by the inversion, and (1.96) projects H^n_+ into D^n, as is geometrically illustrated in [22]. D^n is the classical n-dimensional Lobachevsky space, and D^2 is the well-known Lobachevsky disc.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

The metric (1.93) of $\hat{D}^n$ is mapped into $d\tilde{s}^2 = 4(|\tilde{\mathbf{p}}|^2 - 1)^{-2}|\mathbf{d}\tilde{\mathbf{p}}|^2$, which agrees with (1.93), and $K(D^n) = 1$. The geodesics of D^n are the missing arcs of the geodesics of $\hat{D}^n$. See Figure 1.13.

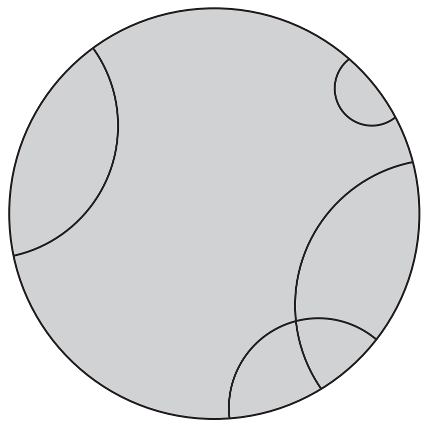

Figure 1.13 Lobachevsky disc and geodesics.

It is very interesting that the geodesics of D^n correspond to the hodograph curves *not* of the Kepler problem with the attractive gravitational force field, but of the central repelling inverse square force field, defining the *central repelling force problem*,

$$\dot{\tilde{\mathbf{q}}} = \tilde{H}_{\tilde{\mathbf{p}}}, \quad \dot{\tilde{\mathbf{p}}} = -\tilde{H}_{\tilde{\mathbf{q}}}, \tag{1.97}$$

$\tilde{\mathbf{q}} = (\tilde{q}_1, \ldots, \tilde{q}_n)$, and

$$\tilde{H} = \frac{1}{2}|\tilde{\mathbf{p}}|^2 + \frac{1}{|\tilde{\mathbf{q}}|}.$$

Collision does not occur for (1.97) for $\tilde{H} = h < \infty$.

Analogous to Theorem 1.32, we have the following theorem and corollary.

Theorem 1.33 *For a positive constant h, the energy surface $\tilde{H}^{-1}(h)$ is topologically equivalent to $T_1(H^n_+)$ embedded in $\mathbf{L^{n+1}}$, and the flow of (1.97) is mapped into the geodesic flow of H^n_+ after a change (1.88) of the independent variables.*

Corollary 1.34 *The hodographs of the central repelling force problem are the geodesics of the classical Lobachevsky space. They are mapped isometrically into the hodographs of the Kepler problem for positive energy. These hodographs represent the geodesics of the inverted Lobachevsky space obtained by inversion with respect to the origin of the classical Lobachevsky space.*

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

More generally, we can state the following theorem [24].

Theorem 1.35 *The Kepler problem for positive energy is mapped into the central repelling problem by inversion with respect to the classical Lobachevsky space.*

Lastly, we mention the case of $h = 0$. The geodesic flow of $\mathbf{R^n}$ is given by $\boldsymbol{\xi'} = \boldsymbol{\eta}, \quad \boldsymbol{\eta'} = \mathbf{0}, \quad ' \equiv \frac{d}{ds}, \quad \boldsymbol{\xi} = (\xi_1, \ldots, \xi_n), \boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$, where $K(\mathbf{R^n}) = 0$. Let $\mathbf{R^n_0} = \mathbf{R^n} \backslash (0, 0, \ldots, 0)$ and $T_1(\mathbf{R^n}) = \{(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbf{R^{2n}} || \boldsymbol{\eta}|^2 = \sum_{k=1}^n \eta_k^2 = 1, < \boldsymbol{\xi}, \boldsymbol{\eta} >= 0\}$, where $<,>$ is the standard Euclidean inner product.

Theorem 1.36 *The energy surface $H^{-1}(0)$ is topologically equivalent to $T^1(\mathbf{R^n_0})$, and the geodesic flow on $\mathbf{R^n}$ is mapped into the Kepler flow after a transformation of time.*

In this case the mapping into $\mathbf{p}$-space is given by

$$p_k = \frac{\xi_k}{|\boldsymbol{\xi}|^2}, \qquad k = 1, 2, \ldots, n \tag{1.98}$$

which is just inversion with respect to $\boldsymbol{\xi} = \mathbf{0}$. This maps the geodesics of $\mathbf{R^n}$, given by $\boldsymbol{\xi''} = \mathbf{0}$, into the families of circles passing through the origin $\mathbf{p} = \mathbf{0}$, which are the hodographs. As in the other cases, (1.98) is canonically extended and time is transformed by (1.88) in order to map the geodesic flow into the Kepler flow, which is regular at collision in the geodesic coordinates, $\boldsymbol{\xi}, \boldsymbol{\eta}, s$.

1.6.3 A Simplified General Proof

A relatively short proof is given to prove the equivalence of the Kepler flow and the flow of the central repelling problem with the geodesic flows on the spaces of constant curvature for all energy cases at once.

Let $Q(a)$ be a family of n-dimensional quadratic manifolds defined by

$$Q(a) = \left\{ \mathbf{x} = (x_0, x_1, \ldots, x_n) \in \mathbf{R^{n+1}} | \sum_{k=1}^n x_k^2 + a^{-3} x_0^2 = a^{-1} \right\}, \tag{1.99}$$

$a \in \mathbf{R^1}$. $Q(a)$ for different a yields the following surfaces:

- For $a = 0$, $Q(a) = \mathbf{R^n}$, with coordinates $\tilde{\mathbf{x}} = (x_1, \ldots, x_n)$.

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

- For $a > 0$, $Q(a)$ is a family of n-dimensional ellipsoids.
- For $a < 0$, $Q(a)$ is a family of hyperboloids of two sheets.
- For $a \gg 0$, the ellipsoids are very thin, centered along the x_0-axis, with large semimajor axis. As a approaches zero, they become more spherical in shape, and for $a = 1$, $Q(a)$ is the n-sphere, S^n. For $0 < a < 1$ the sphere flattens out. As $a \downarrow 0$, $|x_0| \to 0$, and $Q(a)$ becomes infinitely thin, converging to $\mathbf{R^n}$ for $a = 0$. For $a \lesssim 0$, $\mathbf{R^n}$ bifurcates to a flattened two-sheeted hyperboloid with max and min vertice points at $x_0 = \pm a$. When $a = -1$, the standard two-sheeted hyperboloid $H^n_{+,-}$, considered previously, is obtained, and for $a \ll 0$, the two sheets become very thin, lying near the x_0-axis.

Next, embed $Q(a)$ into the respective family of spaces $\mathbf{L^{n+1}}(a)$ defined by the metric

$$d\tilde{s}^2 = \sum_{k=1}^{n} dx_k^2 + a^{-3}dx_0^2, \tag{1.100}$$

which is Riemannian.

For notation, with $\mathbf{y} = (y_0, y_1, \ldots, y_n), \mathbf{z} = (z_0, z_1, \ldots, z_n)$, we set

$$\langle \mathbf{y}, \mathbf{z} \rangle = \sum_{k=1}^{n} y_k z_k + a^{-3} y_0 z_0$$

and

$$\|\mathbf{y}\|^2 = \langle \mathbf{y}, \mathbf{y} \rangle.$$

Thus, $Q(a)$ is given by $\|\mathbf{x}\|^2 = a^{-1}$.

Lemma 1.37 *The geodesics $\mathbf{x} = \mathbf{x}(s)$ on $Q(a) \subset \mathbf{L^{n+1}}(a)$ are given by the system*

$$\mathbf{x}'' + a\mathbf{x} = \mathbf{0}, \tag{1.101}$$

$' \equiv \frac{d}{ds}$ *on the unit tangent bundle of $Q(a)$,*

$$\begin{aligned} T_1(Q(a)) = \{\ &\mathbf{x} \in \mathbf{R^{n+1}}, \mathbf{x}' \in \mathbf{R^{n+1}} | \|\mathbf{x}\|^2 = a^{-1}, \\ &\|\mathbf{x}'\| = 1, \langle \mathbf{x}, \mathbf{x}' \rangle = 0\}. \end{aligned} \tag{1.102}$$

Proof. Set $g(\mathbf{x}) = \|\mathbf{x}\|^2 - a^{-1}$. Thus, $Q(a)$ is given by $g(\mathbf{x}) = 0$. We want to find minimal paths $\mathbf{x} = \mathbf{x}(s)$ on $Q(a), s \in \mathbf{R^1}$. These are given by the *variational problem*

$$\delta \int (f - \mu g) ds = 0 \tag{1.103}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

over all paths on $Q(a)$ connecting any two given points, where $f = \|\mathbf{x}'\|^2$ [128, 180]. The solution to this problem is given by the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial f}{\partial x_k'}\right) - \frac{\partial g}{\partial x_k} = 0, \tag{1.104}$$

$k = 0, 1, \ldots, n$. This is simplified to

$$x_k{}'' - \mu x_k = 0, \tag{1.105}$$

as is verified. We now determine μ. The previous equation implies

$$\langle \mathbf{x}'', \mathbf{x}\rangle - \mu\|\mathbf{x}\|^2 = 0$$

or

$$\langle \mathbf{x}'', \mathbf{x}\rangle = a^{-1}\mu. \tag{1.106}$$

On the other hand,

$$\langle \mathbf{x}', \mathbf{x}\rangle = \frac{1}{2}\frac{d}{ds}\langle \mathbf{x}, \mathbf{x}\rangle = \frac{1}{2}\frac{d}{ds}a^{-1} = 0,$$

and differentiation of this yields

$$\langle \mathbf{x}'', \mathbf{x}\rangle + \|\mathbf{x}'\|^2 = 0.$$

Equation (1.106) then implies $\mu = -a\|\mathbf{x}'\|^2$.

Since $\mathbf{x}(s)$ is a geodesic we know $\|\mathbf{x}'\| =$ constant. We choose $\|\mathbf{x}'\| = 1$, which implies s is arc-length $\tilde{s}$. Thus,

$$\mu = -a. \qquad \square$$

Lemma 1.38 $K(Q(a)) = a, Q(a) \subset \mathbf{L^{n+1}}(a)$.

Proof. We prove this for the case $n = 2$, and general $n > 2$ is an exercise.

A parametric representation of $Q(a)$ is found using geodesic coordinates by using solutions of (1.101),

$$\begin{aligned}\mathbf{x}(s) &= (x_1(s), x_2(s), x_3(s))\\ &= a^{-\frac{1}{2}}(\sin a^{\frac{1}{2}}s, \sin a^{\frac{1}{2}}\beta\cos a^{\frac{1}{2}}s, -a^{\frac{3}{2}}\cos a^{\frac{1}{2}}\beta\cos a^{\frac{1}{2}}s)\end{aligned}$$

and the induced metric $d\tilde{s}^2$ on $Q(a)$ is given by

$$d\tilde{s}^2 = \langle \mathbf{x}_s, \mathbf{x}_s\rangle ds^2 + \langle \mathbf{x}_s, \mathbf{x}_\beta\rangle ds d\beta + \langle \mathbf{x}_\beta, \mathbf{x}_\beta\rangle d\beta^2;$$

$$\begin{aligned}g_{ss} &\equiv \langle \mathbf{x}_s, \mathbf{x}_s\rangle \equiv \langle \mathbf{x}', \mathbf{x}'\rangle = 1,\\ g_{s_\beta} &\equiv \langle \mathbf{x}_s, \mathbf{x}_\beta\rangle = 0,\\ \langle \mathbf{x}_\beta, \mathbf{x}_\beta\rangle &= \cos^2 a^{\frac{1}{2}}s.\end{aligned}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

Thus,

$$g = \begin{vmatrix} g_{ss} & g_{s\beta} \\ g_{\beta s} & g_{\beta\beta} \end{vmatrix} = \cos^2 a^{\frac{1}{2}} s.$$

The Gaussian curvature is given by [128]

$$K = -\frac{1}{2\sqrt{g}}\frac{\partial}{\partial s}\left(\frac{\frac{\partial}{\partial s} g_{\beta\beta}}{\sqrt{g}}\right) = a. \qquad \square$$

For notation, we set $\mathbf{x}^+ = (a, 0, \ldots, 0)$ and

$$Q_0(a) = Q(a)\backslash\{\mathbf{x} = \mathbf{x}^+\}.$$

If $a > 0$, $\mathbf{x}^+$ corresponds to the *north pole* of the corresponding ellipsoid. For $a < 0$, $\mathbf{x}^+$ corresponds to the minimum point of the upper sheet of the two-sheeted hyperboloid.

We consider the Kepler problem (1.80), on the energy surface $H^{-1}(h)$ given by (1.82) with $h = -a/2$.

Theorem 1.39 *The Kepler flow defined by (1.80) on the surface $H^{-1}(-\frac{a}{2})$ is mapped into the geodesic flow of (1.101) on $T_1(Q_0(a))$, where $K(Q_0(a)) = a, Q_0(a) \subset \mathbf{L^{n+1}}(a)$, and $\mathbf{x} = \mathbf{x}^+$ corresponds to collision $\mathbf{q} = 0$, which is regularized by restoring $\mathbf{x}^+$ to $Q_0(a)$. The mapping is given by*

$$p_k = \frac{a^2 x_k}{x_0 - a}, \qquad k = 1, 2, \ldots, n, \tag{1.107}$$

together with the time transformation (1.88).

Proof. Equation (1.88) applied to (1.80) implies

$$\mathbf{p}' = \frac{-\mathbf{q}}{|\mathbf{q}|^2}, \quad \mathbf{q}' = |\mathbf{q}|\mathbf{p}, \quad {}' \equiv \frac{d}{ds}. \tag{1.108}$$

Differentiation of (1.108) yields

$$\mathbf{p}'' = -|\mathbf{q}|^{-1}(\mathbf{p} + 2(\mathbf{p}, \mathbf{q})\mathbf{p}'). \tag{1.109}$$

Some basic identities are needed. Equation (1.107) implies

$$|\mathbf{p}|^2 = \frac{1}{|\tilde{\mathbf{x}}|^2}(x_0 a^{-1} + 1)^2, \tag{1.110}$$

$\tilde{\mathbf{x}} = (x_1, \ldots, x_n)$, and $\mathbf{x} = (x_0, x_1, \ldots, x_n)$. Now, $Q(a)$ can be written as $|\tilde{\mathbf{x}}|^2 = a^{-1}(1 - a^{-1}x_0)(1 + a^{-1}x_0)$, thus (1.110) becomes

$$|\mathbf{p}|^2 = a\frac{a + x_0}{a - x_0}. \tag{1.111}$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

It is noted that (1.111) yields the *inverse map*, using (1.107), which yields

$$x_0 = \frac{|\mathbf{p}|^2 - a}{1 + a^{-1}|\mathbf{p}|^2}, \qquad x_k = -2\frac{a^{-1}p_k}{1 + a^{-1}|\mathbf{p}|^2}, \tag{1.112}$$

$k = 1, 2, \ldots, n$. (1.111) is a map of the norm of $\mathbf{p}$. To obtain a map of $|\mathbf{q}|$, (1.82) implies $|\mathbf{q}|^{-1} = \frac{1}{2}(a+|\mathbf{p}|^2)$, which upon substitution into (1.111) yields

$$|\mathbf{q}| = a^{-2}(a - x_0). \tag{1.113}$$

Finally, differentiation of (1.113) and using (1.108) yields

$$(\mathbf{p}, \mathbf{q}) = -a^{-2}x_0', \tag{1.114}$$

$(\mathbf{p}, \mathbf{q}) = p_1q_1 + \cdots + p_nq_n$.

Equation (1.107) together with the identities (1.113), (1.114) is all we need to transform (1.109).

Differentiating (1.107) twice with respect to s yields

$$p_k' = A_k(x_0 - a)^{-2}, \tag{1.115}$$

$$p_k'' = \frac{(x_0 - 1)(x_k''a^2(x_0 - a) - a^2x_kx_0'') - 2A_kx_0'}{(x_0 - a)^3}, \tag{1.116}$$

where $A_k = (x_0 - a)a^2x_k' - a^2x_kx_0'$. Substituting (1.113), (1.114), (1.107), (1.115) into the right-hand side of (1.109) yields

$$p_k'' = \frac{a^4x_k(x_0 - a) - 2x_0'A}{(x_0 - a)^3}. \tag{1.117}$$

Finally, equating (1.116), (1.117) yields

$$ax_k'' + a^2x_k = x_k''x_0 - x_kx_0'', \tag{1.118}$$

$k = 1, 2, \ldots, n$. This equation is used to conclude the proof by taking inner products. We first show $x_0'' + ax_0 = 0$. Take the inner product of (1.118) with $\tilde{\mathbf{x}}$,

$$a(\tilde{\mathbf{x}}'', \tilde{\mathbf{x}}) + a^2|\tilde{\mathbf{x}}|^2 = (\tilde{\mathbf{x}}'', \tilde{\mathbf{x}})x_0 - |\tilde{\mathbf{x}}|^2x_0''. \tag{1.119}$$

Since $\mathbf{x}, \mathbf{x}' \in T_1(Q_0(a)), |\tilde{\mathbf{x}}|^2 = a^{-1} - a^{-3}x_0^2$ we get an expression for $(\tilde{\mathbf{x}}'', \tilde{\mathbf{x}})$ by differentiating $\langle \mathbf{x}, \mathbf{x}' \rangle = 0$ with respect to s and using $\|\mathbf{x}'\|^2 = 1$, which together yield $(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}'') = -1 - a^{-3}x_0x_0''$. Equation (1.119) becomes

$$(x_0'' + ax_0)(a^{-1}x_0 - 1) = 0.$$

But $\mathbf{x} \neq \mathbf{x}^+$, so that $a^{-1}x_0 - 1 \neq 0$, and we obtain

$$x_0'' + ax_0 = 0.$$

To obtain $\tilde{\mathbf{x}}'' + a\tilde{\mathbf{x}} = \mathbf{0}$, we substitute $x_0'' = -ax_0$ into (1.118). This implies $(\tilde{\mathbf{x}}'' + a\tilde{\mathbf{x}})(a - x_0) = 0$, yielding

$$\tilde{\mathbf{x}}'' + a\tilde{\mathbf{x}} = \mathbf{0}. \qquad \square$$

For general queries, contact webmaster@press.princeton.edu

© Copyright, Princeton University Press. No part of this book may be distributed, posted, or reproduced in any form by digital or mechanical means without prior written permission of the publisher.

The case of the central repelling problem is obtained by considering more generally the Hamiltonian system (1.80) with Hamiltonian

$$H = \frac{1}{2}|\mathbf{p}|^2 - \frac{\Delta}{|\mathbf{q}|} \tag{1.120}$$

on the surface $H^{-1}(-\frac{a}{2}), \Delta \in \mathbf{R^1}$. This case is obtained as a scaling

$$\mathbf{p} \to \Delta^{-\frac{1}{3}}\mathbf{p}, \qquad \mathbf{q} \to \Delta^{-\frac{1}{3}}\mathbf{q}, \qquad \mathbf{a} \to \Delta^{-\frac{2}{3}}\mathbf{a},$$

in all the previous equations in Theorem 1.39 and its proof. Equation (1.101) is replaced by

$$\mathbf{x}'' + a\Delta^{-\frac{2}{3}}\mathbf{x} = 0,$$

$(\mathbf{x}, \mathbf{x}') \in T_1(Q(a\Delta^{-\frac{2}{3}}))$. For example,

$$H = \frac{1}{2}|\mathbf{p}|^2 - \frac{1}{|\mathbf{q}|} = -\frac{a}{2}$$

becomes

$$\tilde{H} = \frac{1}{2}\Delta^{-\frac{2}{3}}|\mathbf{p}| - \frac{\Delta^{\frac{1}{3}}}{|\mathbf{q}|} = -\frac{a}{2}\Delta^{-\frac{2}{3}}$$

and multiplying by $\Delta^{\frac{2}{3}}$ yields (1.120) with $H \equiv \tilde{H}\Delta^{\frac{2}{3}}$. Equation (1.80) is seen to become $\dot{\mathbf{q}} = \mathbf{p}, \dot{\mathbf{p}} = -\Delta\mathbf{q}|\mathbf{q}|^{-3}$.

The results obtained in this section are for the two-body problem. More generally, we could ask if geodesic equivalent flows can be constructed for the $n-$body problem for $n \geq 3$. This was recently answered in an interesting paper by McCord, Meyer, and Offin [149], where, in general the answer is no unless the angular momentum is zero and the energy is positive.

For general queries, contact webmaster@press.princeton.edu