The Kepler problem from a differential geometry point of view

Translation of the original German Version:
Das Keplerproblem differentialgeometrisch gesehen

January 9, 2017

Dr. Thomas S. Ligon
ORCID 0000-0002-4067-876X

Thomas Ligon:
The Kepler problem from a differential geometry point of view

Diploma Thesis, Summer Semester 1973

Advisor: Manfred Schaaf

Faculty of Physics
Theory Department
Ludwig-Maximilians-Universität München

http://nbn-resolving.de/urn:nbn:de:bvb:19-epub-21922-0

The Kepler problem from a differential geometry point of view

Thomas S. Ligon

Abstract

This paper examines the Kepler 2-body problem as an example of the symplectic differential geometric formulation of Hamiltonian mechanics. First, the foundations of symplectic differential geometry and the conventional analysis of the Kepler problem are presented. Then, the SO(4) and SO(3,1) symmetry of the problem and the conserved angular momentum and Runge-Lenz vectors are discussed. The symmetry is also discussed globally, and the integral curves of the Runge-Lenz vector are found.

Physics Faculty of the University of Munich, Theory Department, Chair of Prof. Dr. Fritz Bopp

June, 1973

Contents

Abstract ... i
Symbols ... iii
Introduction ... iv
I. Differential Geometry ... 1
II. Symplectic Differential Geometry ... 6
III. The Significance of the Symplectic Form ... 9
IV. Reduction to a Smaller Dimension ... 11
V. The Kepler Problem and First Reduction ... 12
VI. Stepwise Reduction ... 15
VII. Calculation of the Trajectories of H via a Single Reduction ... 19
VIII. Discussion of the Symmetry and Trajectories of $\boldsymbol{M}$ ... 23
IX. Stereographic Projection and global SO(4) (resp. SO(3,1)) Symmetry ... 28
X. Concluding Remarks ... 34
Literature ... 35

Symbols

Symbols with no definition in the text:

■	end of a proof
[1;(7.12.)]	theorem or definition (7.12.) in literature reference 1
$\otimes$	tensor product
$\boldsymbol{A}$	vector, n-tuple $(A_1,\ldots,A_n)$
$\boldsymbol{A}\cdot\boldsymbol{B}$	$\coloneqq \sum_{i=1}^{n} A_i B_i$
$\lVert\boldsymbol{A}\rVert$	$\coloneqq (\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}$
$\boldsymbol{A}\times\boldsymbol{B}$	$(\boldsymbol{A}\times\boldsymbol{B})_i \coloneqq A_j B_k - A_k B_j$ where (i,j,k) is a cyclic permutation of (1,2,3)

Symbols with definition in the text:

$D(M)$	group of diffeomorphisms of M	(I.4.)
$T_m(M)$	tangent space of M at m	(I.6.)
$T(M)$	tangent bundle of M	(I.6.)
$\tau_M: TM \to M$	tangent bundle projection	(I.6.)
$Tf: TM \to TN$	tangent of f	(I.7.)
$T_s^r(M)$	vector bundle of tensors	(I.9.)
$\mathcal{T}_s^r(M)$	tensor fields	(I.10.)
$\mathcal{F}(M)$	algebra (ring) of functions	(I.10.)
$\mathcal{X}(M)$	vector fields	(I.10.)
$\mathcal{X}^*(M)$	covector fields	(I.10.)
φ^*	induced mapping	(I.11.)
$c'(\lambda)$	tangent of a curve	(I.13.)
F_X	flow of X	(I.16.)
df	differential of f	(I.17.)
$\mathbf{L}_X f$	Lie derivative	(I.17.)
$[X,Y]$	Lie bracket of vector fields	(I.24.)
$\Omega^k(E)$	exterior k-forms on E	(I.28.)
$\mathbf{A}$	alternation mapping	(I.29.)
$\wedge$	exterior product	(I.31.)
$\Omega^k(M)$	algebra of exterior differential forms on M	(I.34.)
$\mathbf{d}$	exterior derivative	(I.35.)
$\mathbf{i}_X$	inner product	(I.36.)
$\flat, \sharp$	vector bundle isomorphisms	(II.1.)
$-\theta_0, \omega_0$	canonical forms on T^*V	(II.5.)
$X_f = (\mathbf{d}f)^\#$	vector field	(II.7.)
$\{f,g\}$	Poisson bracket	(II.7.)
$\mathcal{L}_G$	left-invariant vector fields of G	(II.17.)
ℓ_G	Lie algebra of G	(II.17.)
$\exp(tX)$	exponential mapping	(II.19.)

Introduction

A physical theory has three aspects: 1) a physical scope, 2) a mathematical model, and 3) an interpretation. Here, we consider the scope of non-relativistic motion of material bodies that have been idealized as mass points. As a mathematical model, we choose Hamiltonian dynamical systems on symplectic manifolds (definition: (II.25.)). The (parametrized) integral curves of a Hamiltonian vector field are interpreted as trajectories of the mass points. In traditional text books of mechanics, the name "symplectic manifold" is not mentioned, but its essential properties are already there. As such, classical mechanics can take advantage of a relatively mature mathematical theory, with capabilities that exceed those of the traditional formulation of mechanics.

In recent decades, symplectic differential geometry has gained interest among mathematicians, as discussed in Abraham [1]. Among other things, the global and qualitative properties of dynamic systems are investigated.

However, this abstract mathematical theory is also interesting for physicists, namely in connection with quantum mechanics. The concept of quantization refers to the transition from a classical mechanical description of a problem to a quantum mechanical description. That can be formulated mathematically: Quantization is a functor from the category of symplectic manifolds to the category of complex Hilbert spaces. For the mechanics of mass points, there are rudiments of such a construction (cf. e.g. Kostant [9]), but they don't yet suffice for a comparison with experiments or with conventional methods of quantum mechanics. A loftier goal would be to extend the scope to infinite-dimensional symplectic manifolds, and thus to continuum mechanics and quantum field theory (cf. Segal [14]). Then we could hope to achieve a rigorous formulation of quantum field theory. This way, it might even be possible to solve some of the current problems in quantum field theory.

This quantization technique is also being investigated by mathematicians, since it supplies information about possible unitary representations of Lie groups [9].

Since symplectic differential geometry, which is relevant for physicists, is however an abstract mathematical theory that is not well known, it appeared advisable to calculate a non-trivial, concrete example. This was the motivation for this thesis, in which the Kepler 2-body problem is presented in the language of symplectic differential geometry. Because of that, this thesis contains especially many examples, including some that are not necessary or customary.

One reason for the recent interest in the Kepler problem is the investigation of "non-symmetry" groups. Those are groups that contain the symmetry groups of the problem as subgroups, and contain additional elements that do **not** leave H invariant. This is a situation that is similar to hadron physics: a "broken" symmetry. We attempt to understand this circumstance better on the basis of a known problem (the Kepler problem). Consequently, we want to describe the dynamics on the basis of group theory. One problem is the fact that, in general, it is possible to specify infinitely many groups that break symmetry. Non-symmetry groups were described in terms of quantum mechanics by Bacry [2] and Bander and Itzykson [3] and in terms of classical mechanics by Györgyi [6]. (Györgyi also provides a detailed bibliography.)

In this thesis, we first present the mathematical foundation according to Abraham [1]. This is followed by a discussion of the role played by a symplectic form in mechanics. After the question of the reduction to fewer dimensions has been addressed, the Kepler problem is formulated. The conventional analysis of the problem (conservation of momentum and angular momentum, relative coordinates) is presented following Abraham [1].

The Kepler problem has a pronounced symmetry that has recently been discussed in analogy to hadron physics [6]. This symmetry makes it possible to find the trajectories without solving any differential equations (chapter VII.). Using Poisson brackets, it is easy to find a Lie algebra for this symmetry. But that doesn't tell us if there is also a global action of a group that is associated with the Lie algebra. The answer to this question consists of finding the integral curves of the corresponding vector fields. In chapter VIII., this problem is reduced to a single differential equation and the existence of the integral curves is proven. For non-zero energy, the integral curves are also determined in chapter IX. As far as I know, this is something new.

It is well known that conservation of angular momentum is a consequence of rotational symmetry. But which transformation symmetry leads to conservation of the Runge-Lenz vector (aphelion)? This question is also answered by the determination of the integral curves in chapter IX., but the transformation is not so straightforward as a rotation.

At this point I would like to thank Manfred Schaaf for suggesting this topic, and for many useful discussions.

I. Differential Geometry

In this chapter, some of the most important concepts and theorems of modern differential geometry are presented according to Abraham [1], especially: manifold, tangent bundle, vector field, integral curve, Lie derivative, and exterior derivative. Above all, it is intended to delineate and establish the nomenclature needed to the subsequent material. We don't include proofs. The presentation is also very fragmentary, for example we don't prove that the tangent bundle of a manifold is a manifold itself. We refer to the text books by Abraham [1] and Dieudonné [4].

We would also like to point out that all concepts in this chapter can be defined on the basis of the differentiable structure alone. We will encounter a special structure in the next chapter.

(I.1.) Definition. *[1;(3.1.)] Let S be a set.*

(i) A **local chart** *is a bijection φ between a subset U of S and an open subspace of $\mathbb{R}^n$, where n may depend on φ. An* **atlas** *on S is a family $\mathcal{A}$ of local charts $\{(U_i,\varphi_i)|i\in I\}$ such that*

1) $S=\cup\{U_i|i\in I\}$,

2) For two charts (U_i,φ_i) and (U_j,φ_j) with $U_i\cap U_j\neq\emptyset$, $\varphi(U_i\cap U_j)$ is open in $\mathbb{R}^n$ and $\varphi_{ji}:=\varphi_j\circ\varphi_i^{-1}|\varphi_i(U_i\cap U_j)$ and φ_{ji}^{-1} are bijective and C^∞ (i.e. differentiable arbitrarily often).

(ii) Two atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are **equivalent** *iff $\mathcal{A}_1\cap\mathcal{A}_2$ is an atlas. A* **differentiable structure** *$\mathcal{S}$ on S is an* **equivalence class of atlases** *on S. $\mathcal{A}_\mathcal{S}:=\cup\{\mathcal{A}|\mathcal{A}\in\mathcal{S}\}$ is the* **maximal atlas** *of $\mathcal{S}$, and a local chart $(U,\varphi)\in\mathcal{A}_\mathcal{S}$ is called an* **admissible local chart**.

(iii) A **differentiable manifold** *is a pair $(S,\mathcal{S})$ where S is a set and $\mathcal{S}$ is a differentiable structure on S.*

(I.2.) Definition. *[1;(3.2.)] Let M be a differentiable manifold. A subset $A\subset M$ is* **open** *iff for each $a\in A$ there is an admissible local chart (U,φ) such that $a\in U$ and $U\subset A$.*

A differentiable manifold M is an **n-manifold** *iff for every $a\in M$ there exists an admissible local chart (U,φ) such that $a\in U$ and $\varphi(U)\subset\mathbb{R}^n$. (The dimension is constant.)*

In this thesis, a **manifold** *will always mean a Hausdorff differentiable manifold with a countable base of the topology.*

Remark. The customary definitions of manifold are not completely uniform. It is often required for S to be a topological space and for the local charts to be homeomorphisms. Here, the topology is induced by the local charts, but the result is the same. What we refer to as a maximal atlas or a differentiable structure is often called atlas.

(I.3.) Definition. *[1;(3.5.)] A* **submanifold** *of a manifold M is a subset $B\subset M$ such that for every $b\in B$ there is an admissible local chart (U,φ) with the* **submanifold property**, *i.e.*

$\varphi: U\to\mathbb{R}^k\times\mathbb{R}^l$ and $\varphi(U\cap B)=\varphi(U)\cap(\mathbb{R}^k\times\{0\})$.

(I.4.) Definition. *[1;(3.8.)] A map $f: M\to N$ where M and N are manifolds is called a* **diffeomorphism** *iff f is bijective and f and f^{-1} are of class C^∞. D(M) denotes the group of diffeomorphisms.*

Remark. A diffeomorphism is an isomorphism in the category of differentiable manifolds.

(I.5.) Definition. *[1;(5.1.)] Let M be a manifold and $m\in M$. A* **curve** *at m is a C^1 (i.e. continuously differentiable at least once) map $c: I\to M$ from an open interval $I\subset\mathbb{R}$ with $0\in I$ and $c(0)=m$. Let c1 and c2 be curves at m and (U,φ) an admissible chart with $m\in U$. Then c_1 and c_2 are* **tangent at m with respect to** *φ iff $\varphi\circ c_1$ and $\varphi\circ c_2$ are tangent at 0.*

Remark. (I.5.) defines an equivalence relation which is independent of φ within a differentiable structure. Then $[c]_m$ denotes an equivalence class of curves at m with representative *c*.

(I.6.) Definition. *[1;(5.3.)] Let M be a manifold and* $m \in M$. *The* **tangent space of** M **at** *m is the set of equivalence classes of curves at m.* $T_m(M) := \{[c]_m | c \text{ is a curve at } m\}$. $T(M) := \bigcup_{m \in M} T_m(M)$ *is the* **tangent bundle** *of M.*

The mapping $\tau_M: TM \to M$ *defined by* $\tau_M([c]_m) = m$ *is the* **tangent bundle projection** *of M.*

Remark. $T_m(M)$ is a vector space in a canonical fashion.

(I.7.) Definition. *[1;(5.6.)] Let* $f: M \to N$ *be a* C^1 *mapping. Then* $Tf: TM \to TN$ *defined by* $Tf([c]_m) = [f \circ c]_{f(m)}$ *is the* **tangent** *of f.*

(I.8.) Theorem. *[1;(5.7.)] Let* $\boldsymbol{f}: \boldsymbol{M} \to \boldsymbol{N}$ *and* $\boldsymbol{g}: \boldsymbol{N} \to \boldsymbol{K}$ *be* $\boldsymbol{C^1}$ *mappings of manifolds. Then*

(i) $g \circ f: M \to K$ *is* C^1 *and* $T(g \circ f) = Tg \circ Tf$;

(ii) If $h: M \to M$ *is the identity map, then* $Th: TM \to TM$ *is the identity map;*

(iii) If $f: M \to N$ *is a diffeomorphism, then* $Tf: TM \to TN$ *is a bijection and* $(Tf)^{-1} = T(f^{-1})$.

Remark. In other words, T is a functor. We call it the tangent functor.

(I.9.) Definition. *[1;(6.14.)] Let M be a manifold and* $\tau_M: TM \to M$ *its tangent bundle. The* **vector bundle of tensors of contravariant order** r **and covariant order** s *is*

$$T_s^r(M) = T_s^r(TM) = \bigcup_{m \in M} T_s^r(T_m) = \bigcup_{m \in M} L^{r+s}\left(\underbrace{T_m^*, \dots, T_m^*}_{r\ times}, \underbrace{T_m, \dots, T_m}_{s\ times}, \mathbb{R}\right)$$

where L^{r+s} *is the vector space of multilinear mappings, and* T_m^* *is the dual vector space of* T_m. $T_1^0(M)$ *is called the* ***cotangent bundle*** *and is denoted by* $\tau_M^*: T^*M \to M$.

Remark. For every point $m \in M$, the customary tensor algebra is formed over $T_m(M)$ and then the union is taken over M.

(I.10.) Definition. *[1;(6.15.)] A* **tensor field of type** $\binom{r}{s}$ *on a manifold M is a* C^∞ *section of* $T_s^r(M)$. $\mathcal{T}_s^r(M)$ *denotes the set of all* C^∞ *sections, together with their structure as a real vector space.* $\mathcal{F}(M)$ *denotes the set of* C^∞ *mappings from M in* $\mathbb{R}$ *together with its ring structure:* $(f+g)(x) := f(x) + g(x)$, $(cf)(x) := c(f(x))$, $(fg)(x) := f(x)g(x)$. *A* **vector field** *on M is an element of* $\mathcal{X}(M) := \mathcal{T}_0^1(M)$. *A* **covector field**, *or* **differential 1-form**, *is an element of* $\mathcal{X}^*(M) := \mathcal{T}_1^0(M)$

Remark. A vector field assigns a vector in $T_m(M)$ to every point $m \in M$.

(I.11.) Definition. *[1;(6.16.)] For a diffeomorphism* $\varphi: M \to N$ *and* $t \in \mathcal{T}_s^r(M)$ *define*

$\varphi^* t := (T\varphi)_s^r \circ t \circ \varphi^{-1}$.

(I.12.) Proposition. *[1;(6.17.)]*

(i) $\varphi^* t \in \mathcal{T}_s^r(N)$;

(ii) $\varphi^*: \mathcal{T}_s^r(M) \to \mathcal{T}_s^r(N)$ *is a linear isomorphism.*

Remark. The isomorphism $\varphi: M \to N$ induces an isomorphism $\varphi^*: \mathcal{T}_s^r(M) \to \mathcal{T}_s^r(N)$ in a canonical way.

(I.13.) Definition. *[1;(7.1.)] Let M be a manifold and* $X \in \mathcal{X}(M)$. *An* **integral curve of** X **at** *m is a curve* $c: I \to M$ *at m such that* $X(c(\lambda)) = c'(\lambda) := Tc(\lambda, 1)$ *for each* $\lambda \in I$. *The image of an integral curve of X is the* **trajectory** *of X.*

Remark. $c'(\lambda)$ depends on the parametrization of the curve, not only on the trajectory. If λ denotes time, then $c'(\lambda)$ is the velocity.

(I.14.) Definition. *[1;(7.10.)] Let M be a manifold, X a vector field on M, and* $\mathcal{D}_X \subset M \times \mathbb{R}$ *the set of* $(m, \lambda) \in M \times \mathbb{R}$ *such that there is an integral curve* $c: I \to M$ *of X at m with* $\lambda \in I$. *The vector field is* **complete** *iff* $\mathcal{D}_X = M \times \mathbb{R}$.

Remark. In other words, a vector field is complete if every integral curve can be continued on all of $\mathbb{R}$.

(I.15.) Proposition. *[1;(7.12.)] Let M be a manifold and* $X \in \mathcal{X}(M)$. *Then*

(i) $\mathcal{D}_X \supset M \times \{0\}$;

(ii) $\mathcal{D}_X$ *is open in* $M \times \mathbb{R}$;

(iii) there is a unique mapping $F_X : \mathcal{D}_X \to M$ *such that the mapping* $t \mapsto F_X(m,t)$ *is an integral curve of X at m for all* $m \in M$.

(I.16.) Definition. *[1;(7.13.)] Let M be a manifold and* $X \in \mathcal{X}(M)$. *Then the mapping* F_X *(I.15.) is called the* **integral** *of X, and the curve* $t \mapsto F_X(m,t)$ *is called the* **maximal integral curve** *of X at m. If X is complete,* F_X *is called the* **flow** *of X.*

Remark. "Flow" was called "Fluß" in the original German language thesis, and is called "coulée" in Dieudonné [4;(18.2.2.)]. For an appropriate choice of t, the mapping $m \mapsto F_X(m,t)$ is a diffeomorphism, and if X is complete, $\{F(-,t) | t \in \mathbb{R}\}$ is a one-parameter group of diffeomorphisms. Since $c : \mathbb{R} \to M | t \longmapsto F_X(m,t)$ is an integral curve of X at m, we regain X via differentiation, due to (I.13.). This defines a bijection between complete vector fields and one-parameter groups of diffeomorphisms.

(I.17.) Definition. *[1;(8.1.)] Let* $f \in \mathcal{F}(M)$ *so that*

$Tf : TM \to T\mathbb{R} = \mathbb{R} \times \mathbb{R}$

and

$T_m f = Tf | T_m M \in L(T_m M, \{f(m)\} \times \mathbb{R})$.

We can then define $df : M \to T^*M$ *by* $df(m) = P_2 \circ T_m f$ *where* P_2 *denotes the projection onto the second factor. We call* df *the* **differential** *of* f.

For $X \in \mathcal{X}(M)$, *define* $\boldsymbol{L}_X f : M \to \mathbb{R}$ *by* $\boldsymbol{L}_X f(m) = df(m)(X(m))$. *We call* $\boldsymbol{L}_X f$ *the* **Lie derivative of** f **with respect to** *X.*

Remark. If $\{x_i\}_{i \in I}$ is a system of coordinate functions of M at m, then $\{dx_i\}_{i \in I}$ is a basis of $T_m^*(M)$.

(I.18.) Proposition. *[1;(8.2.)]*

(i) $df \in \mathcal{X}^*(M)$;

(ii) $\boldsymbol{L}_X f \in \mathcal{F}(M)$.

(I.19.) Proposition. *[1;(8.4.)]*

(i) $L_X : \mathcal{F}(M) \to \mathcal{F}(M)$ *is a* **derivation** *on the algebra* $\mathcal{F}(M)$, *i.e.* $\boldsymbol{L}_X$ *is* $\mathbb{R}$ *linear and for* $f, g \in \mathcal{F}(M)$, $\boldsymbol{L}_X(fg) = (\boldsymbol{L}_X(f)g + f\boldsymbol{L}_X(g)$;

(ii) If c is a constant function, then $\boldsymbol{L}_X c = 0$.

(I.20.) Proposition. *[1;(8.5.)] For* $f, g \in \mathcal{F}(M)$ *we have* $d(fg) = (df)g + f(dg)$ *and, if c is constant, dc=0.*

(I.21.) Proposition. *[1;(8.9.)] The collection of operators* $\boldsymbol{L}_X$ *on* $\mathcal{F}(M)$ *forms a real vector space and an* $\mathcal{F}(M)$ *module, with* $(f\mathrm{L}_X)(g) = f(\mathrm{L}_X g)$ *and is isomorphic to* $\mathcal{X}(M)$ *as a real vector space and as an* $\mathcal{F}(M)$ *module. In particular,* $\mathrm{L}_X = 0$ *iff* $X = 0$, *and* $\mathrm{L}_{fX} = f\boldsymbol{L}_X$.

(I.22.) Theorem. *[1;(8.10.)] The collection of all (*$\mathbb{R}$ *linear) derivations on* $\mathcal{F}(M)$ *forms a real vector space isomorphic to* $\mathcal{X}(M)$ *as a real vector space. In particular, for each derivation* $\boldsymbol{\Theta}$ *there is a unique* $X \in \mathcal{X}(M)$ *such that* $\boldsymbol{\Theta} = \boldsymbol{L}_X$.

Remark. Because of this isomorphism, it is possible, and for some applications advantageous, to define the tangent space as the set of derivations of $\mathcal{F}(M)$. (Cf. e.g. Helgason, Differential Geometry and Symmetric Spaces, Academic Press, New York (1962)).

(I.23.) Proposition. *[1;(8.11.)] Let $X, Y \in \mathcal{X}(M)$ be vector fields on M. Then*

$[\boldsymbol{L}_X, \boldsymbol{L}_Y] = \boldsymbol{L}_X \circ \boldsymbol{L}_Y - \boldsymbol{L}_Y \circ \boldsymbol{L}_X$ *is an ($\mathbb{R}$ linear) derivation on $\mathcal{F}(M)$.*

(I.24.) Definition. *[1;(8.12.)] $[X, Y] = \boldsymbol{L}_X Y$ is the unique vector field such that $\boldsymbol{L}_{[X,Y]} = [\boldsymbol{L}_X, \boldsymbol{L}_Y]$. We call $\boldsymbol{L}_X Y$ the* **Lie derivative of** *Y* **with respect to** *X, or the* **Lie bracket of** *X* **and** *Y.*

(I.25.) Proposition. *[1;(8.13.)] The real vector space $\mathcal{X}(M)$, together with the composition $[X, Y]$, forms a* **Lie algebra***. That is,*

(i) $[-,-]$ is $\mathbb{R}$ bilinear;

(ii) $[X, X] = 0$ for all $X \in \mathcal{X}(M)$;

(iii) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathcal{X}(M)$.

(I.26.) Definition. *[1;(8.18.)] If $X \in \mathcal{X}(M)$ we let L_X be the unique differential operator on $\mathcal{T}(M)$ such that it coincides with L_X (I.17.) on $\mathcal{F}(M)$. and with $\boldsymbol{L}_X$ (I.24.) on $\mathcal{X}(M)$. $\boldsymbol{L}_X$ on $\mathcal{T}(M)$ is called the* **Lie derivative with respect to** *X.*

(I.27.) Theorem. *[1;(8.21.)] If $t \in \mathcal{T}(M)$, $\boldsymbol{L}_X t = 0$ iff t is constant along the integral curves of X.*

(I.28.) Definition. *[1;(9.1.)] Let $\boldsymbol{E}$ be a finite dimensional real vector space. Let $\Omega^k(\boldsymbol{E}) = L_a^k(\boldsymbol{E}, \boldsymbol{R})$ be the vector space of skew symmetric k multilinear maps of E in R. An element of $\Omega^k(\boldsymbol{E})$ is called an* **exterior k-form on E***.*

(I.29.) Definition. *[1;(9.2.)] The* **alternation mapping** *$\boldsymbol{A}: T_k^0(\boldsymbol{E}) \to T_k^0(\boldsymbol{E})$ is defined as*

$\boldsymbol{At} := \frac{1}{k!} \sum_{\sigma \in S_k} \varepsilon_\sigma \circ \boldsymbol{t}$

where S_k is the group of permutations of k and ε_σ is the sign of σ.

(I.30.) Proposition. *[1;(9.3.)] $\boldsymbol{A}$ is a linear mapping of $T_k^0(\boldsymbol{E})$ onto $\Omega^k(\boldsymbol{E})$, $\boldsymbol{A}|\Omega^k(\boldsymbol{E})$ is the identity, and $\boldsymbol{A} \circ \boldsymbol{A} = \boldsymbol{A}$.*

(I.31.) Definition. *[1;(9.4.)] If $\boldsymbol{\alpha} \in T_k^0(\boldsymbol{E})$ and $\boldsymbol{\beta} \in T_l^0(\boldsymbol{E})$ define $\boldsymbol{\alpha} \wedge \boldsymbol{\beta} \in \Omega^{k+l}(\boldsymbol{E})$ by $\boldsymbol{\alpha} \wedge \boldsymbol{\beta} := \boldsymbol{A}(\boldsymbol{\alpha} \otimes \boldsymbol{\beta})$. For $\boldsymbol{\alpha} \in T_0^0(\boldsymbol{E}) = \mathbb{R}$, we put $\boldsymbol{\alpha} \wedge \boldsymbol{\beta} = \boldsymbol{\beta} \wedge \boldsymbol{\alpha} = \boldsymbol{\alpha\beta}$.*

(I.32.) Proposition. *[1;(9.5.)] For $\boldsymbol{\alpha} \in T_k^0(\boldsymbol{E})$ and $\boldsymbol{\beta} \in T_l^0(\boldsymbol{E})$, and $\boldsymbol{\gamma} \in T_m^0(\boldsymbol{E})$, we have*

(i) $\boldsymbol{\alpha} \wedge \boldsymbol{\beta} = \boldsymbol{A\alpha} \wedge \boldsymbol{\beta} = \boldsymbol{\alpha} \wedge \boldsymbol{A\beta}$;

(ii) $\wedge$ is bilinear;

(iii) $\boldsymbol{\alpha} \wedge \boldsymbol{\beta} = (-1)^{kl}\, \boldsymbol{\beta} \wedge \boldsymbol{\alpha}$;

(iv) $\boldsymbol{\alpha} \wedge (\boldsymbol{\beta} \wedge \boldsymbol{\gamma}) = (\boldsymbol{\alpha} \wedge \boldsymbol{\beta}) \wedge \boldsymbol{\gamma}$.

(I.33.) Definition. *[1;(9.18.)] We define*

$\omega^k(M) = \omega^k(TM) := \cup_{m \in M} \Omega^k(T_m M)$, $\omega_M^k = \omega^k(\tau_M): \omega^k(M) \to M$ defined by

$\omega^k(\tau_M)(x) = m$ for $x \in \Omega^k(T_m M)$, and

$\omega^k(M)$ is the set of sections of ω_M^k, $\Omega^0(M) = \mathcal{F}(M)$, and $\Omega^1(M) = \mathcal{T}_1^0(M) = \mathcal{X}^*(M)$.

Remark. This way, we have defined the exterior algebra at every $T_m M$ and then taken the union over all m, just as we did for the tensor algebra.

(I.34.) Definition. *[1;(10.3.)] Let $\Omega(M)$ denote the direct sum of $\Omega^k(M)$, $k = 0,1, \dots, n$, together with its structure as an (infinite dimensional) real vector space and with the multiplication $\wedge$ extended component-wise to $\Omega(M)$. $\Omega(M)$ is called the* **algebra of exterior differential forms** *on M. Elements of $\Omega^k(M)$ are called* **k forms***. In particular, elements of $\Omega^1(M) = \mathcal{X}^*(M)$ are called* **1 forms***.*

(I.35.) Theorem. *[1;(10.5.)] Let M be a manifold. Then there is a unique family of mappings*

$\boldsymbol{d}^k(U)\colon \Omega^k(U) \to \Omega^{k+1}(U), k = 0,1,\dots,n$, *and U is open in M, which we denote by* d, *called the* **exterior derivative** *on M, such that*

(i) d *is a* $\wedge$ **antiderivation**. *That is,* d *is* $\mathbb{R}$ *linear and*

for $\alpha \in \Omega^k(U), \beta \in \Omega^l(U), \boldsymbol{d}(\alpha\wedge\beta) = \boldsymbol{d}\alpha\wedge\beta + (-1)^k\alpha\wedge\boldsymbol{d}\beta$;

(ii) If $f \in \mathcal{F}(U), \boldsymbol{d}f = df$ *(I.17.);*

(iii) $\boldsymbol{d} \circ \boldsymbol{d} = 0$ *(that is* $\boldsymbol{d}^{k+1}(U) \circ \boldsymbol{d}^k(U) = 0$*);*

(iv) $\boldsymbol{d}(\alpha|V) = (\boldsymbol{d}\alpha)|V$ *for* $V \subset U$ *open in M.*

(I.36.) Definition. *[1;(10.12.)] Let M be a manifold,* $X \in \mathcal{X}(M)$, *and* $\omega \in \Omega^{k+1}(M)$. *Then we define* $\boldsymbol{i}_X\omega \in \mathcal{T}^0_k(M)$ *by* $\boldsymbol{i}_X\omega(X_1,\dots,X_k) = (k+1)\omega(X, X_1,\dots,X_k)$. *If* $\omega \in \Omega^0(M)$, *we put* $\boldsymbol{i}_X\omega = 0$. *We call* $\boldsymbol{i}_X\omega$ *the* **inner product** *of X and ω.*

(I.37.) Theorem. *[1;(10.13.)] We have* $\boldsymbol{i}_X\colon \Omega^k(M) \to \Omega^{k-1}(M), k = 1,\dots,n$ *and, for* $\alpha \in \Omega^k(M), \beta \in \Omega^l(M), f \in \Omega^0(M)$,

(i) $\boldsymbol{i}_X$ *is a* $\wedge$ **antiderivation***;*

(ii) $\boldsymbol{i}_{fX}\alpha = f\boldsymbol{i}_X\alpha$;

(iii) $\boldsymbol{i}_X\boldsymbol{d}f = \boldsymbol{L}_Xf$;

(iv) $\boldsymbol{L}_X\alpha = \boldsymbol{i}_X\boldsymbol{d}\alpha + \boldsymbol{d}\boldsymbol{i}_X\alpha$;

(v) $\boldsymbol{L}_{fX}\alpha = f\boldsymbol{L}_X\alpha + \boldsymbol{d}f\wedge\boldsymbol{i}_X\alpha$.

(I.38.) Definition. *[1;(10.16.)] We call* $\omega \in \Omega^k(M)$ **closed** *iff* $\boldsymbol{d}\omega = 0$ *and* **exact** *iff there is an* $\alpha \in \Omega^{k-1}(M)$ *such that* $\omega = \boldsymbol{d}\alpha$.

(I.39.) Definition. *[1;(9.14.)] Let* $\boldsymbol{E}$ *and* $\boldsymbol{F}$ *be vector spaces,* $\varphi \in L(\boldsymbol{E},\boldsymbol{F})$, *and* $\boldsymbol{\alpha} \in T^0_k(\boldsymbol{F})$. *Then we define* $\varphi_*\boldsymbol{\alpha} \in L^0_k(\boldsymbol{E})$ *by* $\varphi_*\boldsymbol{\alpha}(\boldsymbol{e}_1,\dots,\boldsymbol{e}_k) = \boldsymbol{\alpha}(\varphi(\boldsymbol{e}_1),\dots,\varphi(\boldsymbol{e}_1))$.

(I.40.) Definition. *[1;(10.7.)] Let M and N be manifolds,* $F\colon M \to N$ *a* C^∞ *mapping, and* $\omega \in \Omega^k(N)$. *Then we define* $F_*\omega\colon M \to \omega^k(M)$ *by* $F_*\omega(m) = (T_mF)_* \circ \omega \circ F(m)$. *(I.39.).*

Remark. This way, a C^∞ mapping of manifolds induces a corresponding mapping in the vector bundle of exterior forms.

II. Symplectic Differential Geometry

In this chapter, we define the concept of a symplectic form, as well as the concepts that depend on it. We define: symplectic manifold, Poisson bracket, Hamiltonian vector field, Hamiltonian function, and Hamiltonian action of a Lie group on a manifold.

(II.1.) Definition. *[1;(14.1.)] Let M be a manifold and* $\omega \in \Omega^2(M)$ *nondegenerate (i.e.* $\omega(m)$ *is a nondegenerate tensor for all* $m \in M$*). Then we define*

$\flat: \mathcal{X}(M) \to \mathcal{X}^*(M) | X \mapsto X^\flat := \boldsymbol{i}_X\, \omega$ *and* $\sharp: \mathcal{X}^*(M) \to \mathcal{X}(M) | \alpha \mapsto \alpha^\sharp | \sharp = \flat^{-1}$.

Remark. ♭ is a vector bundle isomorphism. The situation is similar to a Riemannian metric, but instead of a symmetric tensor we have an asymmetric tensor.

(II.2.) Theorem (Darboux). *[1;(14.7.)] Let ω be a nondegenerate 2-form on a 2n-manifold M. Then* $\boldsymbol{d}\omega = 0$ *iff there is a chart (U,φ) at each* $m \in M$ *such that* $\varphi(m) = 0$ *and with*

$\varphi(u) = (x_1(u), \dots, x_n(u), y_1(u), \dots, y_n(u))$ *we have* $\omega | U = \sum_{i=1}^n dx_i \wedge dy_i$.

(II.3.) Definition. *[1;(14.8.)] A* **symplectic form** *(or a* **symplectic structure***) on a manifold M is a nondegenerate, closed 2-form ω on M. A* **symplectic manifold** *(M,ω) is a manifold M together with a symplectic form ω on M. The charts characterized by the theorem of Darboux are called* **symplectic charts** *and the component functions* x_i, y_i *are called* **canonical coordinates**.

(II.4.) Definition. *[1;(14.9.)] Let (M,ω) and (N,ρ) be symplectic manifolds. A* C^∞ *mapping* $F: M \to N$ *is called* **symplectic** *iff* $F_*\rho = \omega$ *(I.40.).*

(II.5.) Theorem. *[1;(14.14.)] Let V be a manifold and M=T*V. Consider* $\tau_V^*: M \to V$ *and* $T\tau_V^*: TM \to TV$. *For* $v \in V$ *let* $\alpha_v \in M$ *be points in M and* ω_{α_v} *points in TM in the fiber over* α_v. *Define* $\theta_{\alpha_v}: T_{\alpha_v}M \to \mathbb{R} | \omega_{\alpha_v} \mapsto \alpha_v \circ T\tau_V^*(\omega_{\alpha_v})$ *and* $\theta_0: \alpha_v \mapsto \theta_{\alpha_v}$. *Then* $\theta_0 \in \mathcal{X}^*(M)$ *and* $\omega_0 = -\boldsymbol{d}\theta_0$ *is a symplectic form on M.* $-\theta_0$ *and* ω_0 *are called* **canonical forms** *on M.*

Remark. Every manifold has a Riemannian metric, but not necessarily a symplectic form. The theorem above shows that the cotangent bundle *T*V* of every manifold *V* has a symplectic form in a canonical way.

(II.6.) Theorem. *[1;(14.16.)] Let M be a manifold and* $\varphi: M \to M$ *a diffeomorphism. Then* $\varphi^*: T^*M \to T^*M$ *is a symplectic diffeomorphism on* T^*M *with respect to the canonical symplectic structure.*

(II.7.) Definition. *[1;(14.23.)] Let (M,ω) be a symplectic manifold and* $f, g \in \mathcal{F}(M)$. *Then we define* $X_f \in (\boldsymbol{d}f)^\sharp \in \mathcal{X}(M)$, *and the* **Poisson bracket** *of f and g is the function*

$\{f, g\} = -\boldsymbol{i}_{X_f} \boldsymbol{i}_{X_g} \omega$.

(II.8.) Proposition. *[1;(14.24.)] Let (M,ω) be a symplectic manifold and* $f, g \in \mathcal{F}(M)$. *Then* $\{f, g\} = -\boldsymbol{i}_{X_f} \boldsymbol{i}_{X_g} \omega = -\boldsymbol{L}_{X_f} g = +\boldsymbol{L}_{X_g} f$.

(II.9.) Corollary. *[1;(14.25.)] For* $f_0 \in \mathcal{F}(M)$, *the map* $g \mapsto \{f_0, g\}$ *is a derivation on* $\mathcal{F}(M)$.

(II.10.) Proposition. *[1;(14.26.)] Let (M,ω) be a symplectic manifold and* $f, g \in \mathcal{F}(M)$. *Then* $\boldsymbol{d}\{f, g\} = \{\boldsymbol{d}f, \boldsymbol{d}g\} := -\left[\boldsymbol{d}f^\sharp, \boldsymbol{d}g^\sharp\right]^\flat$.

(II.11.) Proposition. *[1;(14.27.)] The real vector space* $\mathcal{F}(M)$, *together with the composition* $\{-,-\}$ *of (II.7.) is a Lie algebra.*

(II.12.) Corollary. *[1;(14.28.)]* $X_{\{f,g\}} = -[X_f, X_g]$.

(II.13.) Proposition. *[1;(14.30.)] Let (M,ω) and (N,ρ) be symplectic manifolds and* $F: M \to N$ *a diffeomorphism. Then F is symplectic iff F preserves the Poisson bracket of all functions, i.e. for all* $f, g \in \mathcal{F}(M)$, $F^*\{f, g\} = \{F^*f, F^*g\}$.

Remark. A symplectic diffeomorphism is an isomorphism in the category of symplectic differentiable manifolds. We recognize the concept of a canonical transformation. (Cf. Goldstein, Classical Mechanics, Addison-Wesley, Reading, Mass. (1950), section 8-3 and 8-4.) A transformation is called canonical when it preserves the form of the Hamilton equations, and that is true exactly when the transformation preserves all Poisson brackets. Investigation of the mathematical model of classical mechanics (in Hamiltonian formalism) is thus synonymous with investigation of the category of symplectic manifolds.

(II.14.) Proposition. *[1;(14.31.)] Let (M,ω) be a 2n-manifold and $f, g \in \mathcal{F}(M)$. Let (U,φ) be a symplectic chart (cf. (II.3.)). Then*

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x_i} \right).$$

(II.15.) Definition. *[1;(16.14.)] Let (M,ω) be a symplectic manifold and $X \in \mathcal{X}(M)$. Then X is called* **globally Hamiltonian** *iff there is an $H \in \mathcal{F}(M)$ such that $X = X_H = (\boldsymbol{d}H)^{\#}$. H is called a* **Hamiltonian function** *for X.*

(II.16.) Proposition. *[1;(16.12.)] Let X be a locally Hamiltonian vector field on a symplectic manifold (M,ω) with a local Hamiltonian function $H \in \mathcal{F}(M)$. Let (V,φ), $V \subset U$, be a symplectic chart with $\varphi(V) \subset \mathbb{R}^{2n}$ and $\varphi(v) = (q_1(v), \dots, q_n(v), p_1(v), \dots, p_n(v))$. Then a curve c(t) on V is an integral curve of X iff*

$$\frac{dq_i}{dt}(c(t)) = \frac{\partial H}{\partial p_i}(c(t)) \text{ for } i=1,\dots,n,$$

and

$$\frac{dp_i}{dt}(c(t)) = -\frac{\partial H}{\partial q_i}(c(t)) \text{ for } i=1,\dots,n.$$

(II.17.) Definition. *[1;(22.1.&22.3.)] A* **Lie group** *G is a manifold together with a group operation, $G \times G \to G | (g_1, g_2) \mapsto g_1 g_2$, which is a smooth mapping of manifolds. The identity element of G is denoted by e.*

A **left translation** *by $g \in G$ is the mapping $L_g: G \to G | g' \mapsto gg'$. A vector field X on G is called* **left invariant** *iff $L_g^* X = X$ for all $g \in G$.*

The set of left invariant vector fields on G forms a Lie subalgebra of $\mathcal{X}(G)$, called the **Lie algebra of left invariant vector fields** *on G and denoted by $\mathcal{L}_G$. The tangent space of G in the point e, $T_e G$, is isomorphic to G as a vector space. $T_e G$, together with the Lie bracket induced by this isomorphism, is called the* **Lie algebra** *of G and is denoted by ℓ_G.*

(II.18.) Proposition. *[1;(22.4.)] Let G be a Lie group and $X \in \mathcal{L}_G$. Then X is complete.*

(II.19.) Definition. *[1;(22.5.)] Let F_X be the flow of $X \in \mathcal{L}_G$. Then $X \mapsto exp(tX) = F_X(e, t)$ is called the* **exponential mapping***.*

(II.20.) Definition. *[1;(22.8.)] Let G be a Lie group and M a manifold. An* **action of** *G* **on** *M is a group homomorphism $\Phi: G \to D(M)$ such that the mapping*

$$ev_\Phi: G \times M \to M | (g, m) \mapsto \Phi(g)(m)$$

is C^∞.

(II.21.) Definition. *[1;(22.9.)] Let $\Phi: G \to D(M)$ be an action of G on M, $x \in \ell_G$ and $X \in \mathcal{L}_G$ with $X(e) = x$. Let $F_X: G \times \mathbb{R} \to G$ be the flow of X. Then*

$$H_X: M \times \mathbb{R} \to M | (m, t) \mapsto \Phi(F_X(e, t))(m) = \Phi(exp(tx))(m)$$

is a flow on M. Let $Y_X \in \mathcal{X}(M)$ be the (unique) vector field such that H_X is the flow of Y_X. Then the Lie algebra homomorphism $\Phi': \ell_G \to \mathcal{X}(M) | x \mapsto Y_x$ is called the **infinitesimal generator** *of Φ and Y_x is called the* **infinitesimal transformation** *of x.*

(II.22.) Definition. *[1;(22.12.)] (i) Let Φ be an action of a Lie group on a symplectic manifold M. Then Φ is called* **Hamiltonian** *iff* $\Phi'(\ell_G) \subset \mathcal{X}_{\mathcal{H}}(M) = \{X \in \mathcal{X}(M) | X \text{ is } Hamiltonian\}$*, i.e. iff every infinitesimal transformation of Φ is globally Hamiltonian.*

(ii) Let $H \in \mathcal{F}(M)$ *and Φ a Hamiltonian action of G such that H is invariant under Φ(g) for all* $g \in G$*. Then G is called a* **symmetry group** *of H under the action Φ.*

(II.23.) Theorem. *[1;(22.13.)] Let* $\Phi: G \to D(M)$ *be a Hamiltonian action on a symplectic manifold M and* $H \in \mathcal{F}(M)$*. Then H is invariant under Φ iff* $\mathbf{L}_Y H = 0$ *for all infinitesimal transformations* $Y \in \Phi'(\ell_G)$*. In this case, if* $K \in \mathcal{F}(M)$ *is a Hamiltonian for Y, or* $Y = X_K = (dk)^\sharp$*, then K is a constant of the motion (i.e. K is constant along the integral curves of* X_H*).*

(II.24.) Definition. *A* **dynamical system** *is a pair (M,X), where M is a differentiable manifold, and* $X \in \mathcal{X}(M)$ *is a vector field on M.*

(II.25.) Definition. *A* **Hamiltonian dynamical system** *is a triple (M,ω,H), where (M,ω) is a symplectic manifold and* $H \in \mathcal{F}(M)$*. Since* $X_H = (dH)^\sharp \in \mathcal{X}(M)$*, every Hamiltonian dynamical system is a dynamical system.*

III. The Significance of the Symplectic Form

In the last chapter, we saw that a symplectic form allows us to define a mapping $\{-,-\}: \mathcal{F}(M) \times \mathcal{F}(M) \to \mathcal{F}(M)$ that creates a Lie algebra structure on $\mathcal{F}(M)$ in such a way that the corresponding derivation $\{-,f\}: \mathcal{F}(M) \to \mathcal{F}(M)$ is the Lie derivative $\mathbf{L}_{X_f}$ on $\mathcal{F}(M)$. This mapping was called the Poisson bracket. Now, in this chapter, we want to consider the opposite direction. We begin with another definition of the Poisson bracket, which was proposed by Dirac [5] and discussed by Pauli [11] and Jost [7]. The chapter ends with a conjecture about the reasons for the importance of the symplectic form in classical mechanics.

(III.1.) Definition. *[7] A* **Poisson bracket** *is a mapping $\{-,-\}: \mathcal{F}(M) \times \mathcal{F}(M) \to \mathcal{F}(M)$ with the following properties: for $f, f_1, f_2, g, h \in \mathcal{F}(M), c \in \mathbb{R}$:*

(a) $\{f, g\} = -\{g, f\} \in \mathcal{F}(M)$,

(b) $\{f_1 + f_2, g\} = \{f_1, g\} + \{f_2, g\}$,

(c) $\{f_1 f_2, g\} = f_1\{f_2, g\} + \{f_1, g\} f_2$,

(d) $\{f, c\} = 0$,

(e) $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$ (Jacobi identity),

(f) the tensor field $\Lambda \in \mathcal{T}_0^2(M)$ defined by $\Lambda(df, dg) = \{f, g\}$ is nowhere degenerate.

Remark. The consistency of Λ in (f) depends on the other properties. The use of the same name of Poisson bracket is justified by the following theorem.

(III.2.) Theorem. *[7]*

(i) A Poisson bracket (III.1.) generates a closed symplectic form ω.

(ii) The Poisson bracket defined by ω (II.7.) and by (III.1.) coincide.

Proof. (i) For the full proof, cf. Jost [7]. Among other things, it shows that the Jacobi identity is fulfilled iff $\mathbf{d}\omega = \mathbf{0}$, i.e. ω is closed. Since Λ is nowhere degenerate, it induces a vector bundle isomorphism $\sharp: \mathcal{X}^*(M) \to \mathcal{X}(M)$, and therefore a nowhere degenerate tensor field $\omega \in \mathcal{T}_0^2(M)$ by means of $\omega(\mathbf{d}f^\sharp, \mathbf{d}g^\sharp) = \{f, g\}$ that is antisymmetric because of (a).

(ii) We have $\omega(\mathbf{d}f^\sharp, \mathbf{d}g^\sharp) = \omega(X_f, X_g) = \mathbf{i}_{X_f}\omega(-, X_g) = \mathbf{i}_{X_g}\mathbf{i}_{X_f}\omega = -\mathbf{i}_{X_f}\mathbf{i}_{X_g}\omega$ (cf. (II.1.), (I.36.), and (II.7.)). ■

(III.3.) Theorem. *A mapping $\{-,-\}: \mathcal{F}(M) \times \mathcal{F}(M) \to \mathcal{F}(M)$ is a Poisson bracket iff*

(i) The vector space $\mathcal{F}(M)$ and the operation $\{-,-\}$ form a real Lie algebra,

(ii) $\{-, g\}$ is a derivation on the associated algebra $\mathcal{F}(M)$, and

(iii) The tensor field $\Lambda \in \mathcal{T}_0^2(M)$ defined by $\Lambda(df, dg) = \{f, g\}$ is nowhere degenerate.

Proof. "(III.1.)⇒(III.3.)(i)" In (c), we take $f_1 \in \mathbb{R}$. Then we have $\{f_1 f_2, g\} = f_1\{f_2, g\} + \{f_1, g\} f_2$, but because of (d) and (a) we have $\{f_1, g\} = 0$, and therefore $\{f_1 f_2, g\} = f_1\{f_2, g\}$. As a result, and because of (b), $\{-,-\}$ is $\mathbb{R}$ linear in the first argument. Because of (a), $\{-,-\}$ is also $\mathbb{R}$ bilinear. In (a), we set $f = g$. Then $\{f, f\} = -\{f, f\}$, and so $\{f, f\} = 0$ for all $f \in \mathcal{F}(M)$. (e) is the Jacobi identity. As a result, the Poisson bracket is a Lie algebra.

"(III.1.)⇒(III.3.)(ii)" As above, $\{-, g\}$ is $\mathbb{R}$ bilinear, and because of (c) it is a derivation on $\mathcal{F}(M)$.

"(III.1.) ⇒(III.3.)(iii)" This is just (f).

"(III.3.)⇒(III.1.)(i)" Because of $\mathbb{R}$ bilinearity, (b) holds. Also, because of the Lie algebra property,

$0 = \{f + g, f + g\} = \{f + g, f\} + \{f + g, g\} = \{f, f\} + \{g, f\} + \{f, g\} + \{g, g\} = \{g, f\} + \{f, g\} \Longrightarrow \{f, g\} = -\{g, f\}$, so (a) holds as well. (c) holds because $\{-, g\}$ is a derivation on $\mathcal{F}(M)$.

Because of $\mathbb{R}$ bilinearity, $\{f_1 f_2, g\} = f_1\{f_2, g\}$ for $f_1 \in \mathbb{R}$. Then, because of the derivation property (c) $\{f_1, g\} = 0$ for $f_1 \in \mathbb{R}$, so (d) holds as well. Since $\{-,-\}$ is a Lie algebra, (e) also holds, and (f) is nothing other than (iii). ■

So we see that the symplectic form is equivalent to the Poisson bracket. In addition, it is known that classical mechanics can be formulated "coordinate free" via Poisson brackets ("equations of motion in Poisson bracket form"). We have also seen that the Poisson bracket is nothing other than a Lie algebra structure and a derivation on the algebra of functions $\mathcal{F}(M)$. The real-valued functions $\mathcal{F}(M)$ also have an important physical meaning: they are exactly the observables in mechanics. In terms of the symplectic form, and in particular via the mapping $(d\,-)^\sharp\colon \mathcal{F}(M) \to \mathcal{X}(M) | f \mapsto (df)^\sharp = X_f$, the observables generate local diffeomorphisms (and sometimes global groups of diffeomorphisms). So we recognize the symplectic form (and the Poisson bracket) as a possibility to represent classical mechanics "coordinate free", whereas the meaning of the observables as generators of transformations is emphasized.

IV. Reduction to a Smaller Dimension

One way to simplify a dynamical system is to reduce its dimension. The first theorem of this chapter provides conditions that make it possible to find a submanifold to which the system can be reduced without any loss of information. The second theorem provides a condition under which the reduction leads to a Hamiltonian dynamic system.

If the functions in theorem (IV.1.) generate a symmetry group (II.22.), then the theorem is an example for using a symmetry to simplify the problem. When the reduced system is Hamiltonian as well, it can be interpreted as "equivalent" to another mechanical problem.

(IV.1.) Theorem. *[4;(16.8.9.)] Let Y be a differentiable manifold, and* $(f_i)_{1\leq i\leq r}$ *a finite series of functions in* $\mathcal{F}(M)$. *Let* $X \coloneqq \{x \in Y | f_i(x) = 0 \text{ for all } 1 \leq i \leq r\}$. *Assume that, for every* $x \in X$, *the differentials* $(d_x f_i)_{1\leq i\leq r}$ *are linearly independent covectors in* $T_x(Y)^*$. *Then*

(i) X is a closed submanifold of Y, and for every $x \in X, T_x(X)$ *is in the kernel of every* $d_x f_i$.

(ii) We have $dim_x(X) = dim_x(Y) - r$.

Remark. This theorem allows us, first without regard for the symplectic form, to reduce the manifold, if we have functions as described above, and $\{f_i, H\} = 0$. Because of (I.27.) and (II.8.), the trajectories of H are contained in the submanifold. The trajectory is thus contained in the surface of constant f_i, and, in particular, because $\{H, H\} = 0$, in the surface of constant energy.

(IV.2.) Theorem. *Let (M,ω) be a symplectic manifold and N a submanifold of M. Then* $(N, \omega|N)$ *is a symplectic manifold iff* $\omega|N$ *is not degenerate.*

Proof. Because of (II.3.), we need to check if $\varphi|N$ is (i) antisymmetrical and (ii) closed. (i) If $\varphi|N$ were not antisymmetrical, there would be $X_1, X_2 \in \mathcal{X}(N)$ with $\omega(X_1, X_2) \neq -\omega(X_2, X_1)$. But, since $\mathcal{X}(N) \subset \mathcal{X}(M)$, ω would not be antisymmetrical, which is a contradiction.

(ii) According to Abraham [1;(10.9.)], d is natural with respect to mappings, i.e. for a C^∞ mapping $F: N \to M$ and $\omega \in \Omega^k(M)$, $F_*\omega \in \Omega^k(N)$ (cf. I.40.) and $F_*(d\omega) = d(F_*\omega)$. For a submanifold $N \subset M$ and $i: N \to M$ the (C^∞) inclusion, we have $\mathrm{i}_*(d\omega) = d(\mathrm{i}_*\omega) \Rightarrow (d\omega)|N = d(\omega|N)$, so $d\omega = 0 \Rightarrow d(\omega|N) = 0$, i.e. ω is closed $\Rightarrow \omega|N$ is closed. ■

V. The Kepler Problem and First Reduction

In this chapter, we formulate the Kepler problem in the language of symplectic differential geometry. Then we perform the customary “reduction to an equivalent 1-body problem”, i.e. we move to relative coordinates.

(V.1.) Technical Lemma. *Let $M = T^*\mathbb{R}^n$ with the canonical symplectic structure (II.5.). Then we choose $q_i, p_i, i = 1, \dots, n$ as canonical coordinates (II.3.). The q_i's are the coordinates of $\mathbb{R}^n$, which exist globally. Since $M = T^*\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$, the p_i's also exist globally. With that, we have*

$\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$.

(i) Let $X \in \mathcal{X}(M)$ with $X: M \to TM | (\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{q}, \boldsymbol{p}; \boldsymbol{x}, \boldsymbol{y})$. Then

$X^\flat = \boldsymbol{x} \cdot d\boldsymbol{p} - \boldsymbol{y} \cdot d\boldsymbol{q}$.

(ii) Let $\alpha \in \mathcal{X}^*(M)$ with $\alpha = \boldsymbol{x}' \cdot d\boldsymbol{q} + \boldsymbol{y}' \cdot d\boldsymbol{p}$. Then

$\alpha^\sharp: M \to TM | (\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{q}, \boldsymbol{p}; \boldsymbol{y}', -\boldsymbol{x}')$.

Proof. (i) According to (II.1.), we have $X^\flat = \mathbf{i}_X\omega = \omega(X, -) = \sum_{i=1}^{n}(dq_i \wedge dp_i)(X, -) = \sum_{i=1}^{n}(dq_i \otimes dp_i - dp_i \otimes dq_i)(X, -) = \sum_{i=1}^{n}(x_i dp_i - y_i dq_i) = \boldsymbol{x} \cdot d\boldsymbol{p} - \boldsymbol{y} \cdot d\boldsymbol{q}$.

(ii) That is just (i) in the reverse direction. ■

(V.2.) Definition. *[1;(32.1.)] Model I for the (2-body) Kepler problem is the system (M, ω, μ, H^μ), with*

(i) $M = T^*W, W = \mathbb{R}^3 \times \mathbb{R}^3 \setminus \Delta, \Delta = \{(\boldsymbol{q}, \boldsymbol{q}) | \boldsymbol{q} \in \mathbb{R}^3\}$ with the canonical symplectic form ω,

(ii) $\mu \in \mathbb{R}, \mu > 0$, and

(iii) $H^\mu \in \mathcal{F}(M)$ defined by

$$H^\mu(\boldsymbol{q}, \boldsymbol{q}', \boldsymbol{p}, \boldsymbol{p}') = \frac{\|\boldsymbol{p}\|^2}{2\mu} + \frac{\|\boldsymbol{p}'\|^2}{2} - \frac{1}{\|\boldsymbol{q} - \boldsymbol{q}'\|}$$

where $\boldsymbol{q}, \boldsymbol{q}' \in \mathbb{R}^3, \boldsymbol{p}, \boldsymbol{p}' \in \mathbb{R}^{3*}$, and $\|-\|$ is the Euclidean norm in $\mathbb{R}^3$.

Remark. Δ is removed from M so that H^μ is in $\mathcal{F}(M)$. X_{H^μ} is not complete for all $m \in M$, since the integral curve can run through Δ (collision).

(V.3.) Proposition (Conservation of linear momentum). *[1;(32.2.)] In Model I, the components of $\boldsymbol{p} + \boldsymbol{p}'$ are constants of the motion.*

Proof. We consider the group $G = (\mathbb{R}^3, +)$ with the action Φ (II.20.) on W, defined by $\Phi(\boldsymbol{r}) | (\boldsymbol{q}, \boldsymbol{q}') \mapsto (\boldsymbol{q} + \boldsymbol{r}, \boldsymbol{q}' + \boldsymbol{r})$. Δ is invariant under Φ. Since Φ is a diffeomorphism, the induced action on M is a symplectic diffeomorphism Φ^* with $\Phi^*(\boldsymbol{r}) = \Phi(\boldsymbol{r})^* | (\boldsymbol{q}, \boldsymbol{q}', \boldsymbol{p}, \boldsymbol{p}') \mapsto (\boldsymbol{q} + \boldsymbol{r}, \boldsymbol{q}' + \boldsymbol{r}, \boldsymbol{p}, \boldsymbol{p}')$. Now we consider a generator of the group $(\boldsymbol{0}, \boldsymbol{r_0}) \in \ell_G$ (as a set, ℓ_G is T_eG, cf. (II.17.)). The corresponding left invariant vector field on G is $X \in \mathcal{L}_G$. $X(\boldsymbol{r}) = (\boldsymbol{r}, \boldsymbol{r_0})$. The flow of X is $F_X(t, \boldsymbol{r}) = \boldsymbol{r} + \mathrm{t}\boldsymbol{r_0}$. The flow on M induced by Φ is $H_X(t, \boldsymbol{q}, \boldsymbol{q}', \boldsymbol{p}, \boldsymbol{p}') = (\boldsymbol{q} + \mathrm{t}\boldsymbol{r_0}, \boldsymbol{q}' + \mathrm{t}\boldsymbol{r_0}, \boldsymbol{p}, \boldsymbol{p}')$. The corresponding vector field (infinitesimal transformation, cf. (II.21.)) is $Y_X(\boldsymbol{q}, \boldsymbol{q}', \boldsymbol{p}, \boldsymbol{p}') = (\boldsymbol{q}, \boldsymbol{q}', \boldsymbol{p}, \boldsymbol{p}', \boldsymbol{r_0}, \boldsymbol{r}'_0, 0,0)$. According to (V.1.) we then have

$$Y_X^\flat = (-\boldsymbol{0}) \cdot d\boldsymbol{q} + (-\boldsymbol{0}) \cdot d\boldsymbol{q}' + (\boldsymbol{r_0}) \cdot d\boldsymbol{p} + (\boldsymbol{r_0}) \cdot d\boldsymbol{p}' = \boldsymbol{r_0} \cdot (d\boldsymbol{p} + d\boldsymbol{p}') = d((\boldsymbol{r_0}) \cdot (\boldsymbol{p} + \boldsymbol{p}')).$$

This gives us $Y_X = (d(\boldsymbol{r_0} \cdot \boldsymbol{p} + \boldsymbol{r_0} \cdot \boldsymbol{p}'))^\sharp$, so that every infinitesimal transformation is Hamiltonian, and thus according to (II.22.) Φ^* is Hamiltonian. H is obviously invariant under Φ^*, and, since $\boldsymbol{r_0} \cdot \boldsymbol{p} + \boldsymbol{r_0} \cdot \boldsymbol{p}'$ is a Hamiltonian function for Y_X, according to (II.23.), $\boldsymbol{r_0} \cdot (\boldsymbol{p} + \boldsymbol{p}')$ is a constant of the motion. As a result, $\boldsymbol{p} + \boldsymbol{p}'$ is a constant of the motion. ■

(V.4.) Definition.

$\widetilde{\boldsymbol{q}} \coloneqq \frac{\mu \boldsymbol{q} + \boldsymbol{q}'}{\mu + 1}$ *(***coordinates relative to the center of mass***)*

$\widetilde{\boldsymbol{q}}' := \boldsymbol{q} - \boldsymbol{q}'$ (**relative coordinates**).

Remark. This gives us

$d\widetilde{\boldsymbol{q}} = \frac{\mu}{\mu+1} d\boldsymbol{q} + \frac{1}{\mu+1} d\boldsymbol{q}'$,

$d\widetilde{\boldsymbol{q}}' = d\boldsymbol{q} - d\boldsymbol{q}'$,

$\boldsymbol{p} = \mu d\boldsymbol{q}$,

$\boldsymbol{p}' = d\boldsymbol{q}'$,

$\widetilde{\boldsymbol{p}} = (\mu + 1) d\widetilde{\boldsymbol{q}}$,

$\widetilde{\boldsymbol{p}}' = \frac{\mu}{\mu+1} d\widetilde{\boldsymbol{q}}'$,

where $\mu + 1$ is the **total mass** and $\frac{\mu}{\mu+1}$ is the **reduced mass**.

(V.5.) Lemma.

$\widetilde{\boldsymbol{p}} = \boldsymbol{p} + \boldsymbol{p}'$ *and is called the* **momentum relative to the center of mass**,

$\widetilde{\boldsymbol{p}}' = \frac{1}{\mu+1}(\boldsymbol{p} - \mu\boldsymbol{p}')$ *and is called the* **relative momentum**,

$\boldsymbol{q} = \widetilde{\boldsymbol{q}} + \frac{1}{\mu+1}\widetilde{\boldsymbol{q}}'$,

$\boldsymbol{q}' = \widetilde{\boldsymbol{q}} - \frac{\mu}{\mu+1}\widetilde{\boldsymbol{q}}'$,

$\boldsymbol{p} = \frac{\mu}{\mu+1}\widetilde{\boldsymbol{p}} + \widetilde{\boldsymbol{p}}'$,

$\boldsymbol{p}' = \frac{1}{\mu+1}\widetilde{\boldsymbol{p}} - \widetilde{\boldsymbol{p}}'$,

$H = \frac{\|\boldsymbol{p}\|^2}{2\mu} + \frac{\|\boldsymbol{p}'\|^2}{2} - \frac{1}{\|\boldsymbol{q}-\boldsymbol{q}'\|} = \frac{\|\widetilde{\boldsymbol{p}}\|^2}{2(\mu+1)} + \frac{\|\widetilde{\boldsymbol{p}}'\|^2}{2\mu/(\mu+1)} - \frac{1}{\|\widetilde{\boldsymbol{q}}'\|}$.

Proof. This is a straightforward calculation. ∎

Remark. Since $\widetilde{\boldsymbol{p}} = \boldsymbol{p} + \boldsymbol{p}'$ is a constant of the motion, H and

$H' := H - \frac{\|\widetilde{\boldsymbol{p}}\|^2}{2(\mu+1)} = \frac{\|\widetilde{\boldsymbol{p}}'\|^2}{2\mu/(\mu+1)} - \frac{1}{\|\widetilde{\boldsymbol{q}}'\|}$

have the same integral curves with respect to $\widetilde{\boldsymbol{p}}'$ and $\widetilde{\boldsymbol{q}}'$. This leads us to Model II (VI.1.).

Remark. The transition from Model I to Model II reduces the dimension of the model by 6. This transition is only partially an example of a reduction due to the three constants of the motion and theorem (IV.1.). In fact, we have a decoupling: the new Hamilton equations are

$\dot{\widetilde{\boldsymbol{q}}} = \frac{\widetilde{\boldsymbol{p}}}{\mu+1}$,

$\dot{\widetilde{\boldsymbol{p}}} = 0$,

$\dot{\widetilde{\boldsymbol{q}}}' = \frac{\widetilde{\boldsymbol{p}}'}{\mu/(\mu+1)}$,

$\dot{\widetilde{\boldsymbol{p}}}' = \frac{\widetilde{\boldsymbol{q}}'}{\|\widetilde{\boldsymbol{q}}'\|^3}$.

As a result, $\widetilde{\boldsymbol{q}}$ and $\widetilde{\boldsymbol{p}}$ are decoupled from $\widetilde{\boldsymbol{q}}'$ and $\widetilde{\boldsymbol{p}}'$. The phase space is

$M = M_1 \times M_2$ with $M_1 \cong \mathbb{R}^3 \times \mathbb{R}^3, M_2 \cong (\mathbb{R}^3 \backslash \{\boldsymbol{0}\}) \times \mathbb{R}^3, H \in \mathcal{F}(M), H = H_1 + H_2$ with

$H_1 = \frac{\|\widetilde{\boldsymbol{p}}\|^2}{2(\mu+1)}$,

$H_2 = \frac{\|\tilde{\boldsymbol{p}}'\|^2}{2\mu/(\mu+1)} - \frac{1}{\|\tilde{\boldsymbol{q}}'\|}.$

Now, H_1 (resp. H_2) is constant on M_2 (resp. M_1). We have thus **reduced** the dimension by three, since $\dot{\tilde{\boldsymbol{p}}} = 0$. In addition, we have **ignored** three dimensions (coordinates relative to the center of mass), since they are trivial ($\tilde{\boldsymbol{q}} = \frac{\tilde{\boldsymbol{p}}t}{\mu+1}$). This reduction of the dimension by 6 results in a **Hamiltonian** dynamic system (Model II), which we can interpret as the **equivalent 1-body problem**.

VI. Stepwise Reduction

We begin by presenting the second model for the Kepler problem; it will also be used in the following chapters as a basis. Model II is equivalent to a 1-body problem and has a 6-dimensional phase space. Based on certain constants of the motion, the problem will be reduced stepwise. The rotational symmetry allows a reduction to 4 dimensions. Due to the conservation of another vector, we can reduce the problem to a 2-dimensional symplectic manifold, Model IV. In this 2-dimensional manifold, we determine the equation for the trajectory.

The reduction based on conservation of angular momentum is customary and is instructive, because it shows how the motion is constrained by the symmetry. In this case, there are three independent constants of the motion, but the problem is only reduced by two dimensions. The final reduction, to the 2-dimensional model IV, is not customary. This reduction, together with theorem (VI.15.), in which we find the equation for the trajectory, serves as a further example of a reduction and of the method of differential geometry.

(VI.1.) Definition. *[1;(32.4.)]* **Model II** *for the Kepler problem is the triple (M,μ,H) with*

(i) $M = T^*U, U = \mathbb{R}^3\backslash\{\mathbf{0}\}$,

(ii) $\mu \in \mathbb{R}, \mu > 0$,

(iii) $H \in \mathcal{F}(M)$ *defined by* $H(\boldsymbol{q},\boldsymbol{p}) = \frac{\|\boldsymbol{p}\|^2}{2\mu} - \frac{1}{\|\boldsymbol{q}\|}$.

Remark. μ now refers to the reduced mass.

(VI.2.) Theorem (Conservation of angular momentum). *[1;(32.5.)] In Model II, the following quantities are constants of the motion:*

$(L_1, L_2, L_3) \coloneqq (q_2p_3 - q_3p_2, q_3p_1 - q_1p_3, q_1p_2 - q_2p_1)$,

with $\boldsymbol{q} = (q_1, q_2, q_3)$ *and* $\boldsymbol{p} = (p_1, p_2, p_3)$.

Proof. We consider the Lie group $G = SO(3)$ of rotations in $\mathbb{R}^3$. The point $\{\mathbf{0}\}$ is invariant, and the action of G on $\mathbb{R}^3\backslash\{\mathbf{0}\}$ is the restriction of the above. This induces a symplectic action on $T^*(\mathbb{R}^3\backslash\{\mathbf{0}\})$ (cf. (II.6.)), that consists of rotations in $\boldsymbol{q}$ space U and the same rotations in $\boldsymbol{p}$ space. We consider for the moment the one-parameter subgroup of rotations about the q_3 axis, with action Φ^*. Then we have, explicitly,

$$\Phi^*|(\boldsymbol{q},\boldsymbol{p}) \mapsto (q_1\cos\theta + q_2\sin\theta, q_2\cos\theta - q_1\sin\theta, q_3, p_1\cos\theta + p_2\sin\theta, p_2\cos\theta - p_1\sin\theta, p_3).$$

The infinitesimal transformation on M is then $Y(\boldsymbol{q},\boldsymbol{p}) = (\boldsymbol{q},\boldsymbol{p}; q_2, -q_1, 0, p_2, -p_1, 0)$. Then we have $Y^\flat = q_1dp_2 + p_2dq_1 - q_2dp_1 - p_1dq_2 = d(q_1p_2 - q_2p_1)$, so we have

$Y = (d(q_1p_2 - q_2p_1))^\sharp$, so Φ^* is Hamiltonian, with Hamilton function $q_1p_2 - q_2p_1 = L_3$. The other components are analogous. ■

(VI.3.) Lemma. *The three covectors* $dL_1, dL_2,$ *and* dL_3 *are linearly independent for* $L \neq 0$.

Proof. $0 = \boldsymbol{c}\cdot d\boldsymbol{L} = (\boldsymbol{c}\times\boldsymbol{q})\cdot d\boldsymbol{p} - (\boldsymbol{c}\times\boldsymbol{p})\cdot d\boldsymbol{q} \Rightarrow \boldsymbol{c}\times\boldsymbol{q}$ and $\boldsymbol{c}\times\boldsymbol{p} = \mathbf{0}$. Since $\boldsymbol{q} \neq 0$, we have three possibilities:

(1) $\boldsymbol{p} = \mathbf{0}, \boldsymbol{c} \parallel \boldsymbol{q} \Rightarrow \boldsymbol{L} = \mathbf{0} \Rightarrow$ collision trajectory.

(2) $\boldsymbol{p} \neq \mathbf{0}, \boldsymbol{c} \parallel \boldsymbol{q}, \boldsymbol{c} \parallel \boldsymbol{p} \Rightarrow \boldsymbol{q} \parallel \boldsymbol{p} \Rightarrow \boldsymbol{L} = \mathbf{0} \Rightarrow$ collision trajectory.

(3) $\boldsymbol{p} \neq \mathbf{0}, \boldsymbol{c} \nparallel \boldsymbol{q} \Rightarrow \boldsymbol{c} = \mathbf{0} \Rightarrow$ linearly independent. ■

(VI.4.) Lemma. *Let* (L_1, L_2, L_3) *be defined as in (VI.2.). Then*

(i) $\{L_i, L_j\} = L_k$ *with* (i,j,k) *a cyclic permutation of* $(1,2,3)$,

(ii) If L_i *and* L_j, *for* $i \neq j$, *are constants of the motion, then* L_k, *for* $i \neq j, j \neq k, k \neq i$ *is also a constant of the motion.*

Proof. (i) According to (II.4.) we have e.g.

$$\{L_1, L_2\} = \sum_{i=1}^{n} \left(\frac{\partial L_1}{\partial q_i} \frac{\partial L_2}{\partial p_i} - \frac{\partial L_1}{\partial p_i} \frac{\partial L_2}{\partial q_i} \right) = q_1 p_2 - q_2 p_1 = L_3$$

(ii) Because of the Jacobi identity (II.11.), we have $\{\{L_1, L_2\}, H\} + \{\{L_2, H\}, L_1\} + \{\{H, L_1\}, L_2\} = 0$. Since L_1 and L_2 are constants of the motion, $\{L_2, H\} = \{H, L_1\} = 0$, and so

$\{L_3, H\} = \{\{L_1, L_2\}, H\} = 0$. ■

Remark. Starting from the three-dimensional group $SO(3)$, we have identified three constants of the motion (in contrast to the remark in Abraham [1], pg 191, stating that $SO(3)$ is two-dimensional.). The fact that the three angular momentum functions are not independent with respect to the Poisson bracket does not disturb the reduction. According to (IV.1.), it is still possible to reduce by three dimensions. However, the reduction by three dimensions cannot lead to a symplectic manifold. (A 2-form on a manifold of uneven dimension would necessarily be degenerate!) That would lead to a loss of all benefits of a symplectic structure, so it appears as inappropriate to reduce by three dimensions. We will now see that a reduction by two dimensions can be useful.

(VI.5.) Lemma. $\boldsymbol{q} \cdot \boldsymbol{L} = \boldsymbol{p} \cdot \boldsymbol{L} = 0$.

Proof. We just need to substitute $\boldsymbol{q} \cdot \boldsymbol{L} = \sum_{i=1}^{3} q_i L_i = 0$, and similarly for $\boldsymbol{p}$.

Remark. This shows that the trajectories of H lie within a plane that is perpendicular to $\boldsymbol{L}$, both in $\boldsymbol{q}$ space and in $\boldsymbol{p}$ space. Since we have seen in (VI.2.) that a rotation in $\boldsymbol{q}$ space and the same rotation in $\boldsymbol{p}$ space is a symplectic diffeomorphism, we can assume without loss of generality that $\boldsymbol{L}$ is oriented in the q_3 (resp. p_3) direction. This leads us to Model III.

(VI.6.) Definition. *[1;(32.6.)]* **Model III** *for the Kepler problem is the triple (M,μ,H) with*

(i) $M = T^*(\mathbb{R}^2 \backslash \{\boldsymbol{0}\})$, *with the canonical symplectic form ω, and* $L = q_1 p_2 - q_2 p_1 \neq 0$,

(ii) $\mu \in \mathbb{R}, \mu > 0$,

(iii) $H \in \mathcal{F}(M)$ *defined by* $H(\boldsymbol{q}, \boldsymbol{p}) = \frac{\|\boldsymbol{p}\|^2}{2\mu} - \frac{1}{\|\boldsymbol{q}\|}$, *where* $\|-\|$ *is the Euclidean norm on* $\mathbb{R}^2$.

Remark. We can interpret Model III as an equivalent 2-dimensional problem.

In order to make a further reduction possible, we make use of a vector $\boldsymbol{A}$, the full meaning of which will become apparent in chapter VIII.

(VI.7.) Definition. *In Model III, we define*

$$L = q_1 p_2 - q_2 p_1$$

$$A_i := \frac{q_i}{\|\boldsymbol{q}\|} + \frac{1}{\mu}[p_i(\boldsymbol{q} \cdot \boldsymbol{p}) - q_i \|\boldsymbol{p}\|^2], i = 1,2$$

(VI.8.) Lemma.

(i) $\{L, H\} = 0, \{A_i, H\} = 0$.

(ii) dA_1 *and* dA_2 *are linearly independent almost everywhere.*

(iii) $dA_1 = \left(\frac{q_2^2}{\|\boldsymbol{q}\|^3} - \frac{p_2^2}{\mu}\right) dq_1 + \left(\frac{-q_1 q_2}{\|\boldsymbol{q}\|^3} + \frac{p_1 p_2}{\mu}\right) dq_2 + \left(\frac{q_2 p_2}{\mu}\right) dp_1 + \left(\frac{q_2 p_1 - 2 q_1 p_2}{\mu}\right) dp_2$,

$dA_2 = \left(\frac{-q_1 q_2}{\|\boldsymbol{q}\|^3} + \frac{p_1 p_2}{\mu}\right) dq_1 + \left(\frac{q_1^2}{\|\boldsymbol{q}\|^3} - \frac{p_1^2}{\mu}\right) dq_2 + \left(\frac{q_1 p_2 - 2 q_2 p_1}{\mu}\right) dp_1 + \left(\frac{q_1 p_1}{\mu}\right) dp_2$.

Note. There is a typographical error in the original version, where the sign in the second term for dA_1 in (iii) is wrong. We have corrected it here (more details in the supplementary information).

Proof. The claims follow directly from the corresponding definitions.

(VI.9.) Definition. $\boldsymbol{A_0} \coloneqq \boldsymbol{A}(m)$, *where* $m \in M$ *is the corresponding initial condition.*

Remark. Because of (IV.1.), $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ is a 2-dimensional submanifold of M. We want to see if $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ is also symplectic. Because of (IV.2.), we only need to check if $\omega|\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ is degenerate.

(VI.10.) Lemma. *Let X be a vector field on* $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ *with* $X|(\boldsymbol{q},\boldsymbol{p}) \mapsto (\boldsymbol{q},\boldsymbol{p},x_1,x_2,x_3,x_4)$. *Then*

$x_3 = \gamma_{31}x_1 + \gamma_{32}x_2$, *and*

$x_4 = \gamma_{41}x_1 + \gamma_{42}x_2$, *with*

$\gamma_{31} = \frac{\mu}{L}\left(\frac{q_1q_2}{\|\vec{q}\|^3} - \frac{p_1p_2}{2\mu}\right)$,

$\gamma_{32} = \frac{\mu}{L}\left(\frac{-q_1^2}{\|\vec{q}\|^3} + \frac{p_1^2}{2\mu}\right)$,

$\gamma_{41} = \frac{\mu}{L}\left(\frac{q_2^2}{\|\vec{q}\|^3} - \frac{p_2^2}{2\mu}\right)$,

$\gamma_{42} = \frac{\mu}{L}\left(\frac{-q_1q_2}{\|\vec{q}\|^3} + \frac{p_1p_2}{2\mu}\right)$.

Proof. Since X is a vector field on $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$, it is in the kernel of dA_1 and dA_2 (cf. (IV.2.)), i.e. $dA_1(X) = dA_2(X) = 0$. The result follows, after a long calculation, from this and formula (VI.8.) (iii). ■

(VI.11.) Lemma. $\omega|\boldsymbol{A}^{-1}(\boldsymbol{A_0}) = \left(\frac{\mu H}{L}\right) dq_1 \wedge dq_2$, *so it is nondegenerate for* $H \neq 0$.

Proof. Let X,Y be vector fields on $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$. Then we have

$\omega(X,Y) = x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2$,

$\omega(X,Y) = x_1(\gamma_{31}y_1 + \gamma_{32}y_2) + x_2(\gamma_{41}y_1 + \gamma_{42}y_2) - y_1(\gamma_{31}x_1 + \gamma_{32}x_2) - y_2(\gamma_{41}x_1 + \gamma_{42}x_2)$

$\omega(X,Y) = (\gamma_{41} - \gamma_{32})(x_2y_1 - x_1y_2)$,

$\omega(X,Y) = \left(\frac{-\mu H}{L}\right)(x_2y_1 - x_1y_2) \Rightarrow$

$\omega|\boldsymbol{A}^{-1}(\boldsymbol{A_0}) = \left(\frac{\mu H}{L}\right) dq_1 \wedge dq_2$. ■

(VI.12.) Definition. **Model IV** *for the Kepler problem is the pair (*$\boldsymbol{A}^{-1}(\boldsymbol{A_0})$*,H') with*

(i) $\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ *is the symplectic submanifold of M in Model III (VI.6.) defined by (VI.7.) and (VI.9.) and symplectic form* $\omega|\boldsymbol{A}^{-1}(\boldsymbol{A_0})$.

(ii) $H' \in \mathcal{F}(\boldsymbol{A}^{-1}(\boldsymbol{A_0})), H' \coloneqq H|\boldsymbol{A}^{-1}(\boldsymbol{A_0})$ *with H defined in (VI.6.).*

(VI.13.) Lemma. $H' = \frac{-(\|\boldsymbol{A}\|^2-1)}{2(q_1A_1+q_2A_2-\|\boldsymbol{q}\|)}$.

Proof. $(q_1A_1 + q_2A_2 - \|\boldsymbol{q}\|) = \frac{1}{\mu}[(\boldsymbol{q}\cdot\boldsymbol{p})^2 - \|\boldsymbol{q}\|^2\|\boldsymbol{p}\|^2] = \frac{-\|\boldsymbol{L}\|^2}{\mu}$,

$\|\boldsymbol{A}\|^2 - 1 = \frac{2H\|\boldsymbol{L}\|^2}{\mu} \Rightarrow H = \frac{\mu(\|\boldsymbol{A}\|^2-1)}{2\|\boldsymbol{L}\|^2} = \frac{-(\|\boldsymbol{A}\|^2-1)}{2(q_1A_1+q_2A_2-\|\boldsymbol{q}\|)}$.

(VI.14.) Definition.

$\cos\varphi \coloneqq \frac{q_1A_1+q_2A_2}{\|\boldsymbol{q}\|\|\boldsymbol{A}\|}$,

$\sin\varphi \coloneqq \frac{q_2A_1-q_1A_2}{\|\boldsymbol{q}\|\|\boldsymbol{A}\|}$.

(VI.15.) Theorem.

(i) $\dot{q} = \left(\frac{\|\boldsymbol{L}\|}{\mu\|\boldsymbol{q}\|}\right)\left(\frac{-\|\boldsymbol{A}\|\, sin\, \varphi}{1-\|\boldsymbol{A}\|\, cos\, \varphi}\right)$,

(ii) $\dot{\varphi} = \frac{\|\boldsymbol{L}\|}{\mu\|\boldsymbol{q}\|^2}$,

(iii) $\|\boldsymbol{q}\| = \frac{const.}{1-\|\boldsymbol{A}\|\, cos\, \varphi}$.

Proof. $\frac{\partial H}{\partial q_2} = \frac{+(\|\mathbf{A}\|^2-1)(A_2-q_2/\|\mathbf{q}\|)}{2(q_1A_1+q_2A_2-\|\mathbf{q}\|)^2} = \frac{-(A_2-q_2/\|\mathbf{q}\|)H}{(q_1A_1+q_2A_2-\|\mathbf{q}\|)}$,

$\dot{q}_1 = \{q_1, H\} = \left(\frac{\|\mathbf{L}\|}{\mu H}\right)\frac{\partial H}{\partial q_2} = \left(\frac{-\|\mathbf{L}\|}{\mu}\right)\frac{(A_2-q_2/\|\mathbf{q}\|)}{(q_1A_1+q_2A_2-\|\mathbf{q}\|)}$,

$\dot{q}_2 = \{q_2, H\} = \left(\frac{\|\mathbf{L}\|}{\mu H}\right)\frac{\partial H}{\partial q_1} = \left(\frac{+\|\mathbf{L}\|}{\mu}\right)\frac{(A_1-q_1/\|\mathbf{q}\|)}{(q_1A_1+q_2A_2-\|\mathbf{q}\|}$,

$\dot{\|\mathbf{q}\|} = \frac{q_1\dot{q}_1+q_2\dot{q}_2}{\|\mathbf{q}\|} = \left(\frac{\|\mathbf{L}\|}{\mu\|\mathbf{q}\|}\right)\left(\frac{-\|\mathbf{A}\|\sin\varphi}{1-\|\mathbf{A}\|\cos\varphi}\right)$,

$(\dot{\sin}\,\varphi) = \cos\varphi\,\dot{\varphi} = \frac{\dot{q}_1A_2-\dot{q}_2A_1}{\|\mathbf{q}\|\|\mathbf{A}\|} - \frac{(q_1A_2+q_2A_1)\dot{q}}{\|\mathbf{q}\|^2\|\mathbf{A}\|} = \left(\frac{\|\mathbf{L}\|}{\mu\|\mathbf{q}\|^2}\right)\cos\varphi \Rightarrow \dot{\varphi} = \frac{\|\mathbf{L}\|}{\mu\|\mathbf{q}\|^2}$.

$\frac{d\|\mathbf{q}\|}{d\varphi} = \frac{\dot{\|\mathbf{q}\|}}{\dot{\varphi}} = \frac{-\|\mathbf{q}\|\|\mathbf{A}\|\sin\varphi}{1-\|\mathbf{A}\|\cos\varphi} \Rightarrow \frac{d\|\mathbf{q}\|}{\|\mathbf{q}\|} = \frac{+\|\mathbf{A}\|d(\cos\varphi)}{1-\|\mathbf{A}\|\cos\varphi}$,

$\ln\|\mathbf{q}\| = -\ln(1-\|\mathbf{A}\|\cos\varphi) + const. \Rightarrow \|\mathbf{q}\| = \frac{const.}{1-\|\mathbf{A}\|\cos\varphi}$. ■

Remark. Theorem (VI.15.) (iii) is an equation for the trajectory of H in $\boldsymbol{q}$ space. The trajectory in the whole phase space of Model II follows from the definitions of Modell IV and Model III. We will discuss the trajectories of H and the integral curves of H in more detail in the next chapter, where they are found using a different method.

VII. Calculation of the Trajectories of *H* via a Single Reduction

This chapter is based on Model II (VI.1.). We have already shown that the angular momentum vector $\boldsymbol{L}$ is a constant of the motion (VI.2.). However, there is another vector, the Runge-Lenz vector $\boldsymbol{A}$, that is invariant under the action of H. Because of $\boldsymbol{L} \cdot \boldsymbol{M} = 0$, not all six components of these two vectors are independent. In fact, the rank of the mapping $\psi: M \to \mathbb{R}^6 | (\boldsymbol{q}, \boldsymbol{p}) \mapsto (\boldsymbol{L}, \boldsymbol{A})$ is equal to 5, as can be seen from (IV.1.) and (VII.4.). This means that it is possible to reduce the 6-dimensional problem to one dimension without solving a differential equation. The trajectories of H are thus determined by purely algebraic means. One additional step provides the integral curves of H, i.e. the time dependence.

(VII.1.) Definition. *[2,3,6]*

(i) $\boldsymbol{A} = \frac{\boldsymbol{q}}{\|\boldsymbol{q}\|} - \frac{\boldsymbol{p}\times\boldsymbol{L}}{\mu}$. *$\boldsymbol{A}$ is called the Runge-Lenz vector.*

(ii) For $H \neq 0$: $\boldsymbol{M} := \sqrt{\frac{\mu}{2|H|}}\boldsymbol{A}$.

(VII.2.) Proposition.

$A_i = \frac{q_i}{\|\boldsymbol{q}\|} + \frac{p_i(\boldsymbol{q}\cdot\boldsymbol{p}) - q_i\|\boldsymbol{p}\|^2}{\mu}$,

$\|\boldsymbol{A}\|^2 = 1 + \frac{2H}{\mu}\|\boldsymbol{L}\|^2$,

$\boldsymbol{L} \cdot \boldsymbol{A} = \boldsymbol{L} \cdot \boldsymbol{M} = 0$,

$\|\boldsymbol{M}\|^2 + \|\boldsymbol{L}\|^2 = \frac{-\mu}{2H}$ *for* $H < 0$,

$\|\boldsymbol{M}\|^2 - \|\boldsymbol{L}\|^2 = \frac{+\mu}{2H}$ *for* $H > 0$,

$\{H, L_i\} = \{H, A_i\} = \{H, M_i\} = 0$.

Proof. The claims follow directly from the corresponding definitions. ■

(VII.3.) Definition.

(i) $\theta := cos^{-1}\left(\frac{\boldsymbol{q}\cdot\boldsymbol{A}}{\|\boldsymbol{q}\|\cdot\|\boldsymbol{A}\|}\right)$, *with*

(ii) $\boldsymbol{q} \cdot \boldsymbol{p} = -(\|\boldsymbol{q}\|^2\|\boldsymbol{p}\|^2 - \|\boldsymbol{L}\|^2)^{1/2}$,

(iii) $q_{\|} := \frac{\boldsymbol{q}\cdot\boldsymbol{A}}{\|\boldsymbol{A}\|}$,

(iv) $q_{\perp} := \frac{\boldsymbol{q}\cdot(\boldsymbol{L}\times\boldsymbol{A})}{\|\boldsymbol{L}\|\cdot\|\boldsymbol{A}\|}$,

(v) $p_{\|} := \frac{\boldsymbol{p}\cdot\boldsymbol{A}}{\|\boldsymbol{A}\|}$,

(vi) $p_{\perp} := \frac{\boldsymbol{p}\cdot(\boldsymbol{L}\times\boldsymbol{A})}{\|\boldsymbol{L}\|\cdot\|\boldsymbol{A}\|}$.

Remark. θ is the angle between $\boldsymbol{q}$ and $\boldsymbol{A}$ in configuration space. Because of

$\|\boldsymbol{L}\|^2 = \boldsymbol{L} \cdot \boldsymbol{L} = (\boldsymbol{q}\times\boldsymbol{p}) \cdot (\boldsymbol{q}\times\boldsymbol{p}) = \|\boldsymbol{q}\|^2\|\boldsymbol{p}\|^2 - (\boldsymbol{q} \cdot \boldsymbol{p})^2$,

the angle between $\boldsymbol{p}$ and $\boldsymbol{A}$ is determined up to the sign, which is defined by (ii). Because of $\boldsymbol{L} \cdot \boldsymbol{A} = 0$, we have an orthonormal basis: $\frac{\boldsymbol{L}}{\|\boldsymbol{L}\|}, \frac{\boldsymbol{A}}{\|\boldsymbol{A}\|}, \frac{\boldsymbol{L}\times\boldsymbol{A}}{\|\boldsymbol{L}\|\cdot\|\boldsymbol{A}\|}$, and because of $\boldsymbol{q} \cdot \boldsymbol{L} = \boldsymbol{p} \cdot \boldsymbol{L} = 0$, we have $\boldsymbol{q} = (0)\frac{\boldsymbol{L}}{\|\boldsymbol{L}\|} + (q_{\|})\frac{\boldsymbol{A}}{\|\boldsymbol{A}\|} + (q_{\perp})\frac{\boldsymbol{L}\times\boldsymbol{A}}{\|\boldsymbol{L}\|\cdot\|\boldsymbol{A}\|}$.

(VII.4.) Theorem.

(i) $\|\boldsymbol{q}\| = \frac{\|\boldsymbol{L}\|^2}{\mu(1-\|\boldsymbol{A}\|\cos\theta)}$,

(ii) $\boldsymbol{q}\cdot\boldsymbol{p} = \frac{-\|\boldsymbol{L}\|\|\boldsymbol{A}\|\sin\theta}{(1-\|\boldsymbol{A}\|\cos\theta)}$,

(iii) $q_{\|} := \frac{\|\boldsymbol{L}\|^2}{\mu}\frac{\cos\theta}{(1-\|\boldsymbol{A}\|\cos\theta)}$,

(iv) $q_{\perp} := \frac{\|\boldsymbol{L}\|^2}{\mu}\frac{\sin\theta}{(1-\|\boldsymbol{A}\|\cos\theta)}$,

(v) $p_{\|} := \frac{-\mu}{\|\boldsymbol{L}\|}\sin\theta$,

(vi) $p_{\perp} := \frac{-\mu}{\|\boldsymbol{L}\|}(\|\boldsymbol{A}\| - \cos\theta)$.

Proof. The derivation is a direct calculation from the definitions. For example, let's calculate $\|\boldsymbol{q}\|$.

$$\boldsymbol{q}\cdot\boldsymbol{A} = \|\boldsymbol{q}\| - \frac{1}{\mu}\boldsymbol{q}\cdot(\boldsymbol{p}\times\boldsymbol{L}) = \|\boldsymbol{q}\| - \frac{1}{\mu}\boldsymbol{L}\cdot(\boldsymbol{q}\times\boldsymbol{p}) = \|\boldsymbol{q}\| - \frac{\|\boldsymbol{L}\|^2}{\mu} \Rightarrow \cos\theta = \frac{1}{\|\boldsymbol{A}\|} - \frac{\|\boldsymbol{L}\|^2}{\mu\|\boldsymbol{A}\|\|\boldsymbol{q}\|} \Rightarrow$$
$$1 - \|\boldsymbol{A}\|\cos\theta = \frac{\|\boldsymbol{L}\|^2}{\mu\|\boldsymbol{q}\|} \Rightarrow \|\boldsymbol{q}\| = \frac{\|\boldsymbol{L}\|^2}{\mu(1-\|\boldsymbol{A}\|\cos\theta)}.\ \blacksquare$$

Remark. From the Poisson bracket relationships in (VII.2.), we can deduce that L_i and A_i (resp. L_i and M_i) are constants of the motion for $H = 0$ (resp. for $H \neq 0$). With that, we have specified $\boldsymbol{q}$ and $\boldsymbol{p}$ in (VII.4.) in terms of the parameter θ, and therefore found a 1-dimensional submanifold that must contain the trajectory of H, and is thus identical with it. The trajectory is also equal to $(\boldsymbol{L},\boldsymbol{A})^{-1}(\boldsymbol{L_0},\boldsymbol{A_0})$ according to (IV.1.). From the trajectory equations in (VII.4.), we can see that the trajectories, projected on the configuration space ($\boldsymbol{q}$ space), are ellipses (resp. parabolas resp. hyperbolas) for $H < 0$ (resp. $H = 0$ resp. $H > 0$). In $\boldsymbol{p}$ space, the trajectories are circles (resp. arcs) for $H < 0$ (resp. for $H \geq 0$). This also shows us the geometrical meaning of the Runge-Lenz vector $\boldsymbol{A}$: $\boldsymbol{A}$ points in the direction of the aphelion and has a length equal to the numerical excentricity.

In order to determine the integral curves of H, we need to find the correct parametrization, i.e. we need to express the angle θ in terms of the time t.

The derivation of (VII.4.) does not depend on the fact that $\boldsymbol{L}$ and $\boldsymbol{A}$ are constants. The formulas for $\boldsymbol{q}$ and $\boldsymbol{p}$ are still correct, when $\boldsymbol{L}$ and $\boldsymbol{A}$ are not constants, for example along the trajectories of M_1, which we will investigate in the next chapter.

(VII.5.) Lemma. *Let $c: \mathbb{R} \to M | t \mapsto (\boldsymbol{q},\boldsymbol{p})$ be an integral curve of H and θ defined as in (VII.3.). Then*

(i) $\frac{d\theta}{dt} = \frac{\mu}{\|\boldsymbol{L}\|^3}(1 - \|\boldsymbol{A}\|\cos\theta)^2$,

(ii) $t = \frac{\|\boldsymbol{L}\|^3}{\mu}\int\frac{d\theta}{(1-\|\boldsymbol{A}\|\cos\theta)^2}$.

Proof. For $q_{\|}$, we have the Hamilton equation:

$$\frac{dq_{\|}}{dt} = \{q_{\|}, H\} = \frac{\partial H}{\partial p_{\|}} = \frac{p_{\|}}{\mu} \Rightarrow \frac{dq_{\|}}{d\theta}\frac{d\theta}{dt} = \frac{p_{\|}}{\mu}(\theta(t)) \Rightarrow \left(\frac{\|\boldsymbol{L}\|^2}{\mu}\right)\frac{d}{d\theta}\left(\frac{\cos\theta}{1-\|\boldsymbol{A}\|\cos\theta}\right)\left(\frac{d\theta}{dt}\right) = \left(\frac{-1}{\|\boldsymbol{L}\|}\right)\sin\theta.$$

$$\frac{d}{d\theta}\left(\frac{\cos\theta}{1-\|\boldsymbol{A}\|\cos\theta}\right) = \frac{-\sin\theta}{(1-\|\boldsymbol{A}\|\cos\theta)^2} \Rightarrow \left(\frac{\|\boldsymbol{L}\|^2}{\mu}\right)\left(\frac{-\sin\theta}{(1-\|\boldsymbol{A}\|\cos\theta)^2}\right)\left(\frac{d\theta}{dt}\right) = \left(\frac{-1}{\|\boldsymbol{L}\|}\right)\sin\theta \Rightarrow$$

$$\frac{d\theta}{dt} = \frac{\mu}{\|\boldsymbol{L}\|^3}(1 - \|\boldsymbol{A}\|\cos\theta)^2.$$

The other cases are analogous. $\blacksquare$

Remark. (VII.5.) (ii) can be integrated directly (cf. Ryshik and Gradstein [13], page 105,106), but we will abstain from that, since the results are a bit unmanageable. Instead, we will define another angle.

(VII.6.) Definition.

$sin\, u := \frac{\sqrt{1-\|\boldsymbol{A}\|^2}\, sin\, \theta}{1-\|\boldsymbol{A}\|\, cos\, \theta},$

$cos\, u := \frac{-\|\boldsymbol{A}\| + cos\, \theta}{1-\|\boldsymbol{A}\|\, cos\, \theta}.$

Remark. u is called the **excentric anomaly** and is the angle between $\boldsymbol{A}$ and the intersection of the ordinate of $\boldsymbol{q}$ with a circle drawn around the ellipse:

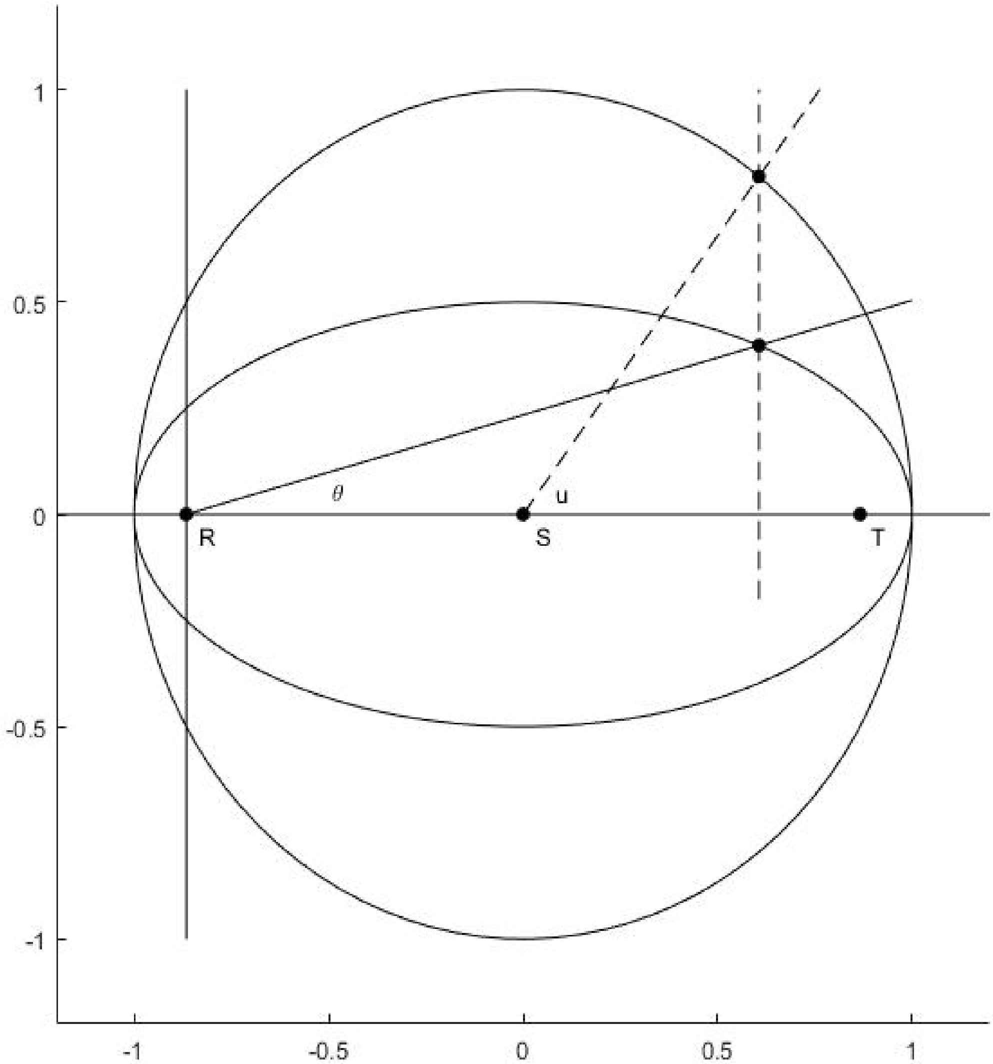

Figure 1. *[1;Figure 33.2] S is the center, R and T are the foci, θ is the angle between* $\boldsymbol{q}$ *and* $\boldsymbol{A}$ *in configuration space (VII.3.), and u is called the* **excentric anomaly** *(VII.6.).*

(VII.7.) Lemma.

$cos\, \theta = \frac{\|\boldsymbol{A}\| + cos\, u}{1+\|\boldsymbol{A}\|\, cos\, u},$

$sin\, \theta = \frac{\sqrt{1-\|\boldsymbol{A}\|^2}\, sin\, u}{1+\|\boldsymbol{A}\|\, cos\, u},$

$1 - \|\boldsymbol{A}\|\, cos\, u = \frac{(1-\|\boldsymbol{A}\|^2)}{1+\|\boldsymbol{A}\|\, cos\, u}.$

Proof. This follows directly from (VII.6.). ■

(VII.8.) Theorem. *Let* $c: \mathbb{R} \to M | t \mapsto (\boldsymbol{q}, \boldsymbol{p})$ *be an integral curve of H and u defined as in (VII.6.) with* $u = 0$ *for* $t = 0$. *Then* $t = \frac{\|\boldsymbol{L}\|^3}{(1-\|\boldsymbol{A}\|^2)^{3/2}}(u + \|\boldsymbol{A}\| \sin u)$.

Proof. $\cos\theta = \frac{\|\mathbf{A}\| + \cos u}{1 + \|\mathbf{A}\| \cos u} \Rightarrow$

$(*)\ \frac{d(\cos\theta)}{dt} = \frac{-(1-\|\mathbf{A}\|^2)\sin u\dot{u}}{(1+\|\mathbf{A}\|\cos u)^2}$.

$\frac{d(\cos\theta)}{dt} = -\sin\theta \frac{d\theta}{dt} \Rightarrow$

$\frac{d(\cos\theta)}{dt} = -\frac{\mu}{\|\boldsymbol{L}\|^3}\sin\theta\,(1 - \|\mathbf{A}\|\cos\theta)^2 = -\frac{\mu}{\|\boldsymbol{L}\|^3}\left(\frac{\sqrt{1-\|\mathbf{A}\|^2}\sin u}{1+\|\mathbf{A}\|\cos u}\right)\left(\frac{(1-\|\mathbf{A}\|^2)}{1+\|\mathbf{A}\|\cos u}\right)^2 \Rightarrow$

$(**)\ \frac{d(\cos\theta)}{dt} = -\frac{\mu}{\|\boldsymbol{L}\|^3}\frac{(1-\|\mathbf{A}\|^2)^{5/2}\sin u}{(1+\|\mathbf{A}\|\cos u)^3}$.

$(*) = (**) \Rightarrow \dot{u} = \frac{du}{dt} = \frac{\mu}{\|\boldsymbol{L}\|^3}\frac{(1-\|\mathbf{A}\|^2)^{3/2}}{1+\|\mathbf{A}\|\cos u} \Rightarrow$

$dt = \frac{\|\boldsymbol{L}\|^3}{\mu(1-\|\mathbf{A}\|^2)^{3/2}}(1 + \|\mathbf{A}\|\cos u)du = \frac{\|\boldsymbol{L}\|^3}{\mu(1-\|\mathbf{A}\|^2)^{3/2}}d(u + \|\mathbf{A}\|\sin u) \Rightarrow$

$\mathrm{t} = \frac{\|\boldsymbol{L}\|^3}{\mu(1-\|\mathbf{A}\|^2)^{3/2}}(u + \|\mathbf{A}\|\sin u)$. ■

VIII. Discussion of the Symmetry and Trajectories of $\boldsymbol{M}$

In the last chapter, we saw that there are two vectors, $\boldsymbol{L}$ and $\boldsymbol{A}$, that are constants of the motion. It is well-known that conservation of $\boldsymbol{L}$ is related to rotational symmetry (cf. (VI.2.)). The six components of $\boldsymbol{L}$ and $\boldsymbol{A}$ (resp. $\boldsymbol{L}$ and $\boldsymbol{M}$) for $H = 0$ (resp. $H \neq 0$) form a closed Lie algebra. Now the question arises: Is there a corresponding global, Hamiltonian action of a group? And if so, how can we describe the action that is related to the conservation of $\boldsymbol{A}$ resp. $\boldsymbol{M}$? For this purpose, we will search for the integral curves of A_1 resp. M_1.

The differential equations are complicated, and we didn't find a direct solution to them. Because of that, we attempted to make use of symmetry. Some functions of $\boldsymbol{L}$ and $\boldsymbol{M}$ that are constant along the integral curves of M_1 are easy to find (VIII.5.). However, among them, only four are independent. If there were five independent and constant functions of $\boldsymbol{L}$ and $\boldsymbol{M}$, then, because of $\boldsymbol{L} \cdot \boldsymbol{M} = 0$, all six components of $\boldsymbol{L}$ and $\boldsymbol{M}$ would be constant. However, that not the case, since $\{M_1, M_2\} = M_3 \neq 0$. Other constants were not found. So it wasn't possible, as in the previous chapter, to solve the problem via reduction alone. A reduction to a smaller dimension as in (VIII.2.) seems to make the problem more complicated.

However, it is possible (cf. (VIII.6.)), to express all components of $\boldsymbol{L}$ and $\boldsymbol{M}$ as functions of the integral curve parameter s. Because of (VII.4.), it is then possible to express the integral curves of M_1 as functions of θ and s. The dependence $\theta(s)$ is not known here, but the problem is reduced to a single, ordinary differential equation of first order (VIII.8.). That proves the existence of the integral curves of A_1 for all H (VIII.9.). Since the integral curves are expressed as a function of s and θ, but $\theta(s)$ is unknown, it is not possible to specify the trajectories here either.

(VIII.1.) Theorem.

(i) For all H:

$\{q_1, A_1\} = (1/\mu)(q_2p_2 + q_3p_3)$,

$\{q_2, A_1\} = (1/\mu)(q_2p_1 - 2q_1p_2)$,

$\{q_3, A_1\} = (1/\mu)(q_3p_1 - 2q_1p_3)$,

$\{p_1, A_1\} = (1/\|\boldsymbol{q}\|^3)(-q_2{}^2 - q_3{}^2) + (1/\mu)(p_2{}^2 + p_3{}^2)$,

$\{p_2, A_1\} = (1/\|\boldsymbol{q}\|^3)(q_1q_2) + (1/\mu)(-p_1p_2)$,

$\{p_3, A_1\} = (1/\|\boldsymbol{q}\|^3)(q_1q_3) + (1/\mu)(-p_1p_3)$,

(ii) For $H \neq 0$:

$$\{q_1, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left(\frac{A_1p_1}{2\mu} - H\left(\frac{1}{\mu}\right)(q_2p_2 + q_3p_3)\right),$$

$$\{q_2, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left(\frac{A_1p_2}{2\mu} - H\left(\frac{1}{\mu}\right)(q_2p_1 - 2q_1p_2)\right),$$

$$\{q_3, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left(\frac{A_1p_3}{2\mu} - H\left(\frac{1}{\mu}\right)(q_3p_1 - 2q_1p_3)\right),$$

$$\{p_1, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left[\frac{A_1}{2}(-\frac{q_1}{\|\boldsymbol{q}\|^3}) - H\left(\frac{-q_2{}^2-q_3{}^2}{\|\boldsymbol{q}\|^3} + \left(\frac{1}{\mu}\right)(p_2{}^2 + p_3{}^2)\right)\right],$$

$$\{p_2, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left[\frac{A_1}{2}(-\frac{q_2}{\|\boldsymbol{q}\|^3}) - H\left(\frac{q_1q_2}{\|\boldsymbol{q}\|^3} + \left(\frac{1}{\mu}\right)(-p_1p_2)\right)\right],$$

$$\{p_3, M_1\} = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left[\frac{A_1}{2}(-\frac{q_3}{\|\boldsymbol{q}\|^3}) - H\left(\frac{q_1q_3}{\|\boldsymbol{q}\|^3} + \left(\frac{1}{\mu}\right)(-p_1p_3)\right)\right],$$

with + sign for $H < 0$ *and – sign for* $H > 0$.

Proof. For $H < 0$:

$$\{x_j, M_i\} = \left\{x_j, \sqrt{\frac{\mu}{2}}(-H)^{-1/2}A_i\right\} = \sqrt{\frac{\mu}{2}}A_i\left(-\frac{1}{2}\right)(-H)^{-3/2}(-1)\{x_j, H\} + \sqrt{\frac{\mu}{2}}(-H)^{-1/2}\{x_j, A_i\}$$

$$\{x_j, M_i\} = \sqrt{\frac{\mu}{2}}(-H)^{-3/2}\left(\frac{A_i}{2}\{x_j, H\} - H\{x_j, A_i\}\right),$$

and similarly for $H > 0$. Substitution of $\{x_j, H\}$ and $\{x_j, A_i\}$ from (II.14.) yields the result. ■

Remark. Because of the properties of the Poisson bracket, we have

$q_i' = \left(\frac{dq_i}{ds}\right) = \{q_i, M_1\} = \mathbf{L}_{X_{M_1}} q_i$ (cf. (II.16.) and (II.8.)). As a result, these are the differential equations for an integral curve of M_1 and of A_1.

(VIII.2.) Lemma. *For $H \neq 0$, the integral curves $c: \mathbb{R} \to M | s \mapsto (\boldsymbol{q}, \boldsymbol{p})$ of M_1 are determined by the following differential equations:*

$$q_1' = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left[\frac{p_1}{\mu}\left(\frac{A_1}{2} + Hq_1\right) - \frac{H}{p_1}\left(\frac{q_1}{\|\boldsymbol{q}\|} + A_1 + 2Hq_1\right)\right],$$

$$p_1' = \pm\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left[\left(\frac{1}{\|\boldsymbol{q}\|}\right)^3\left(\frac{-A_1 q_1}{2} - H{q_1}^2\right) - H\left(\frac{1}{\|\boldsymbol{q}\|} + 2H - \left(\frac{{p_1}^2}{\mu}\right)\right)\right],$$

$$\left(\frac{1}{\|\boldsymbol{q}\|}\right)' = \mp\sqrt{\frac{\mu}{2}}|H|^{-3/2}\left(\frac{1}{\|\boldsymbol{q}\|}\right)^3\left[\left(\frac{A_1}{2} + Hq_1\right)\left(\frac{1}{p_1}\right)\left(\frac{q_1}{\|\boldsymbol{q}\|} + A_1 + 2Hq_1\right) + \left(\frac{-Hp_1\|\boldsymbol{q}\|^2}{\mu}\right)\right].$$

Proof. We can introduce six new coordinates: $q_1, p_1, \left(\frac{1}{\|\mathbf{q}\|}\right), L_1, H,$ and A_1. These six functions are independent, and three of them are constant, namely $L_1, H,$ and A_1. As a result, we need to find the differential equations for $q_1, p_1,$ and $\left(\frac{1}{\|\mathbf{q}\|}\right)$, as functions of these six new coordinates. That follows directly, after a long calculation, from the definitions for $L_1, H,$ and A_1 and from (VIII.1.). ■

Remark. This lemma serves only as an example, and is not used in the following material.

(VIII.3.) Theorem. *Let H, A_i, M_i be defined as in (VI.1.) and (VIII.1.). Then:*

(i) For all H and (i,j,k) = cyclic permutation of (1,2,3):

$\{H, L_i\} = 0,$

$\{L_i, L_j\} = L_k,$

$\{H, A_i\} = 0,$

$\{L_i, A_j\} = A_k,$

$\{L_i, A_i\} = 0,$

$\{A_i, A_j\} = (-2H/\mu)L_k,$

(ii) For $H < 0$:

$\{M_i, M_j\} = L_k,$

$\left\{\left(\frac{L\pm M}{2}\right)_i, \left(\frac{L\pm M}{2}\right)_j\right\} = \left(\frac{L\pm M}{2}\right)_k,$

$\left\{\left(\frac{L+M}{2}\right)_i, \left(\frac{L-M}{2}\right)_j\right\} = 0.$

(iii) For $H > 0$:

$\{M_i, M_j\} = -L_k,$

(iv) For $H = 0$*:*

$\{A_i, A_j\} = 0$.

Proof. It follows directly from the definitions. ■

(VIII.4.) Corollary. *For* $H \neq 0$ *(resp.* $H = 0$*),* $\mathbf{L}$ *and* $\mathbf{M}$ *(resp.* $\mathbf{L}$ *and* $\mathbf{A}$*) form the following Lie algebras:*

(i) For $H < 0$*: Lie algebra of* $O(4)$ *and* $SU(2)\times SU(2)$*,*

(ii) For $H > 0$*: Lie algebra of* $O(3,1)$*,*

(iii) For $H = 0$*: Lie algebra of* $IO(3) = \mathbb{R}^3 \ltimes O(3) =$ *Euclidean group in three dimensions.*

(VIII.5.) Theorem. *For* $H \neq 0$ *(resp.* $H = 0$*), let* $c: \mathbb{R} \to M$ *be an integral curve of* M_1 *(resp.* A_1*). Then the following functions are constant along c:*

(i) For all H:

$L_1, A_1, H, L_2M_2 + L_3M_3$,

(ii) For $H < 0$*:*

$L_2^2 + M_3^2, L_3^2 + M_2^2, M_2M_3 - L_2L_3$,

(iii) For $H > 0$*:*

$L_2^2 - M_3^2, L_3^2 - M_2^2, M_2M_3 + L_2L_3$,

(iv) For $H = 0$*:*

A_1, A_2, A_3, L_1, H.

Proof. Because of (I.27.) and (II.8.), we need to calculate the Poisson brackets, for example in case (ii):

$\{M_2M_3 - L_2L_3, M_1\} = M_2\{M_3, M_1\} + M_3\{M_2, M_1\} - L_2\{L_3, M_1\} - L_3\{L_2, M_1\} =$

$M_2(L_2) + M_3(-L_3) - L_2(M_2) - L_3(-M_3) = 0$. ■

(VIII.6.) Theorem. *For* $H \neq 0$ *(resp.* $H = 0$*), let* $c: \mathbb{R} \to M | s \mapsto (\boldsymbol{q}, \boldsymbol{p})$ *be an integral curve of* M_1 *(resp.* A_1*). Then*

(i) For $H < 0$*:*

$L_2 = +C_1 \cos(s - \alpha_1)$,

$L_3 = +C_2 \cos(s - \alpha_2)$,

$M_2 = -C_2 \sin(s - \alpha_2)$,

$M_3 = +C_1 \sin(s - \alpha_1)$,

$(L\times M)_1 = (C_1^2/2)\sin 2(s - \alpha_1) + (C_2^2/2)\sin 2(s - \alpha_2)$,

(ii) For $H > 0$*:*

$L_2 = +C_3 \cosh(s - \alpha_3)$,

$L_3 = +C_4 \cosh(s - \alpha_4)$,

$M_2 = +C_4 \sinh(s - \alpha_4)$,

$M_3 = -C_3 \sinh(s - \alpha_3)$,

$(L\times M)_1 = (-C_3^2/2)\sinh 2(s - \alpha_3) - (C_4^2/2)\sinh 2(s - \alpha_4)$,

(iii) For $H = 0$*:*

$L_2 = -A_3 s + C_5$,

$L_3 = +A_2 s + C_6$,

$(L\times A)_1 = -(A_2^2 + A_3^2)s - A_2C_6 + A_3C_5$,

where $C_1, \ldots, C_6, \alpha_1, \ldots, \alpha_4 \in \mathbb{R}$ *are the integration constants.*

Note. There is a typographical error in the original version, where the sign in the expressions for $(L\times M)_1$ is wrong. We have corrected it here (more details in the supplementary information).

Proof. (i) Because of (II.8.), we have $\frac{dL_2}{ds} = \mathbf{L}_{X_{M_1}} L_2 = \{L_2, M_1\} = -M_3$, and so

$L_2' = -M_3$ and $M_3' = +L_2 \Rightarrow L_2 + iM_3 = C_1 e^{i(s-\alpha_1)} \Rightarrow$

$L_2 = C_1 \cos(s-\alpha_1), M_3 = +C_1 \sin(s-\alpha_1)$.

(ii) Here we have:

$L_2' = -M_3$ and $M_3' = -L_2 \Rightarrow L_2 + M_3 = C_3 e^{-(s-\alpha_3)}, L_2 - M_3 = C_3 e^{+(s-\alpha_3)}$,

(iii) Since $A_3' = \{A_3, A_1\} = 0$, A_3 is constant. $L_2' = \{L_2, A_1\} = -A_3 \Rightarrow L_2 = -A_3 s + C_5$. ∎

(VIII.7.) Lemma. *With the prerequisites of (VIII.6.), we have:*

(i) For $H < 0$*:*

$$\frac{d\|\boldsymbol{M}\|}{ds} = \frac{-(M_2L_3 - M_3L_2)}{\|\boldsymbol{M}\|} = \frac{+(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{M}\|},$$

$$\frac{d\|\boldsymbol{L}\|}{ds} = \frac{+(M_2L_3 - M_3L_2)}{\|\boldsymbol{L}\|} = \frac{-(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|},$$

(ii) For $H > 0$*:*

$$\frac{d\|\boldsymbol{M}\|}{ds} = \frac{-(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{M}\|},$$

$$\frac{d\|\boldsymbol{L}\|}{ds} = \frac{-(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|},$$

(iii) For $H = 0$*:*

$$\frac{d\|\boldsymbol{L}\|}{ds} = \frac{-(\boldsymbol{L}\times\boldsymbol{A})_1}{\|\boldsymbol{L}\|}.$$

Proof. For example, for $H < 0$: $\frac{d\|\mathbf{M}\|}{ds} = \frac{1}{2\|\mathbf{M}\|}\frac{d\|\mathbf{M}\|^2}{ds} = \frac{1}{2\|\mathbf{M}\|}\{\|\mathbf{M}\|^2, M_1\} =$

$(1/2\|\mathbf{M}\|)(\{M_1^2, M_1\} + \{M_2^2, M_1\} + \{M_3^2, M_1\}) =$

$(1/2\|\mathbf{M}\|)(2M_2\{M_2, M_1\} + 2M_3\{M_3, M_1\}) = (1/\|\mathbf{M}\|)(-M_2L_3 + M_3L_2)$. ∎

(VIII.8.) Theorem. *For* $H \neq 0$ *(resp.* $H = 0$*), let* $c\colon \mathbb{R} \to M | s \mapsto (\boldsymbol{q}, \boldsymbol{p})$ *be an integral curve of* M_1 *(resp.* A_1*) and* θ *defined as in (VII.3.). Then:*

(i) For $H \neq 0$*:*

$$\frac{d\theta}{ds} = \left[\frac{-\mu^2 M_1}{4H|H|\|\boldsymbol{M}\|^2\|\boldsymbol{L}\|^3}\right] + \cos\theta\left[\sqrt{\frac{\mu}{2|H|}}\frac{\mu M_1}{H\|\boldsymbol{M}\|\|\boldsymbol{L}\|^3}\right] + \cos^2\theta\left[\frac{-\mu M_1}{2H\|\boldsymbol{L}\|^3}\right] + \sin\theta\left[-\sqrt{\frac{\mu}{2|H|}}\frac{2(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|^2}\right] + \sin\theta\cos\theta\left[\frac{+(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|^2}\right].$$

(ii) For $H = 0$*:*

$$\frac{d\theta}{ds} = (A_1/\|\boldsymbol{L}\|)(1-\cos\theta)^2 + (1/\|\boldsymbol{L}\|^2)(\boldsymbol{L}\times\boldsymbol{A})_1 \sin\theta(2-\cos\theta).$$

Proof. Because of (VII.4.), we can express $\boldsymbol{q}$ and $\boldsymbol{p}$ as functions of $\boldsymbol{M}$, $\boldsymbol{L}$, and θ. Because of (VIII.6.), $\boldsymbol{M}$ and $\boldsymbol{L}$ can be expressed as functions of s. Then the differential equation for $\theta(s)$ follows from (VIII.1.). The proof is analogous to (VII.5.). Here, we use the equation for $p_{\|}$, because it is the simplest. We sketch the proof for $H \neq 0$*:*

$\frac{dp_\parallel}{ds} = \frac{d}{ds}\left(\frac{-\mu}{\|\boldsymbol{L}\|} sin\theta\right) = \frac{+\mu}{\|\boldsymbol{L}\|^2}\frac{d\|\boldsymbol{L}\|}{ds} sin\theta - \frac{\mu}{\|\boldsymbol{L}\|} cos\theta \frac{d\theta}{ds}$.

With (VIII.7.), we then have:

$(*)\frac{dp_\parallel}{ds} = \left(-\frac{\mu}{L^3}\right)(\boldsymbol{L}\times\boldsymbol{M})_1 sin\theta - \frac{\mu}{\|\boldsymbol{L}\|} cos\theta \frac{d\theta}{ds}$.

On the other hand, we have

$\frac{dp_\parallel}{ds} = \frac{d}{ds}\left(\frac{\boldsymbol{p}\cdot\boldsymbol{M}}{\|\boldsymbol{M}\|}\right) = \frac{d\boldsymbol{p}}{ds}\cdot\frac{\boldsymbol{M}}{\|\boldsymbol{M}\|} + \frac{\boldsymbol{p}}{\|\boldsymbol{M}\|}\cdot\frac{d\boldsymbol{M}}{ds} - \frac{\boldsymbol{p}\cdot\boldsymbol{M}}{\|\boldsymbol{M}\|^2}\frac{d\boldsymbol{M}}{ds}$, and

$\frac{\boldsymbol{p}}{\|\boldsymbol{M}\|}\cdot\frac{d\boldsymbol{M}}{ds} = \frac{\boldsymbol{p}}{\|\boldsymbol{M}\|}\cdot\{\boldsymbol{M}, M_1\} = \left(\frac{1}{\|\boldsymbol{M}\|}\right)(\mp(\boldsymbol{p}\times\boldsymbol{L})_1) = \pm\left(A_1 - \frac{q_1}{\|\boldsymbol{q}\|}\right)\left(\frac{\mu}{\|\boldsymbol{M}\|}\right)$, with the upper sign for $H < 0$ and the lower sign for $H > 0$. This implies

$\frac{dp_\parallel}{ds} = \frac{d\vec{p}}{ds}\cdot\frac{\vec{M}}{\|\boldsymbol{M}\|} \pm \left(A_1 - \frac{q_1}{\|\boldsymbol{q}\|}\right)\left(\frac{\mu}{\|\boldsymbol{M}\|}\right) \mp \frac{p_\parallel}{\|\boldsymbol{M}\|^2}(\boldsymbol{L}\times\boldsymbol{M})_1$. From (VIII.1.) we get:

$\frac{d\vec{p}}{ds}\cdot\frac{\vec{M}}{\|\boldsymbol{M}\|} = \frac{q_\parallel M_1}{2H\|\boldsymbol{q}\|^3} + \sqrt{\frac{\mu}{2|H|}}\left(\frac{-M_1}{\|\boldsymbol{M}\|\|\boldsymbol{q}\|} + \frac{q_1 q_\parallel}{\|\boldsymbol{q}\|^3} + \frac{\|\boldsymbol{p}\|^2 M_1}{\mu\|\boldsymbol{M}\|} - \frac{p_1 p_\parallel}{\mu}\right) \Rightarrow$

$\frac{dp_\parallel}{ds} = \mp\frac{p_\parallel}{\|\boldsymbol{M}\|^2}(\boldsymbol{L}\times\boldsymbol{M})_1 \mp \frac{\mu q_1}{\|\boldsymbol{M}\|\|\boldsymbol{q}\|} + \frac{q_\parallel M_1}{2H\|\boldsymbol{q}\|^3} + \sqrt{\frac{\mu}{2|H|}}\left(\frac{+M_1}{\|\boldsymbol{M}\|\|\boldsymbol{q}\|} + \frac{q_1 q_\parallel}{\|\boldsymbol{q}\|} - \frac{p_1 p_\parallel}{\mu}\right)$.

Now, we use (VII.4.) and (VII.2.) and the expression

$q_1 = q_\parallel\left(\frac{1}{\|\boldsymbol{M}\|}\right)M_1 + q_\perp\left(\frac{1}{\|\boldsymbol{L}\|\|\boldsymbol{M}\|}\right)(\boldsymbol{L}\times\boldsymbol{M})_1$, and arrive at

$(**)\frac{dp_\parallel}{ds} = cos\theta\left[\mp\frac{\mu^3 M_1}{4H^2\|\boldsymbol{M}\|^2\|\boldsymbol{L}\|^4}\right] + cos^2\theta\left[-\sqrt{\frac{\mu}{2|H|}}\frac{\mu^2 M_1}{H\|\boldsymbol{M}\|\|\boldsymbol{L}\|^4}\right] + cos^3\theta\left[\frac{\mu^2 M_1}{2H\|\boldsymbol{L}\|^4}\right] +$
$sin\theta cos\theta\left[\sqrt{\frac{\mu}{2|H|}}\frac{2\mu(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|^3\|\boldsymbol{M}\|}\right] + sin\theta cos^2\theta\left[\frac{-\mu(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|^3}\right] + sin\theta\left[\frac{-\mu(\boldsymbol{L}\times\boldsymbol{M})_1}{\|\boldsymbol{L}\|^3}\right]$.

A comparison of (*) and (**) yields the result. ■

(VIII.9.) Theorem. *For $H \neq 0$ (resp. $H = 0$) let $L_1 \neq 0 \neq M_1$ (resp. $L_1 \neq 0$). Then*

(i) For all $s_0, \theta_0 \in \mathbb{R}$, there is an integral curve of (VIII.8.) (i) and (VIII.8.) (ii) with $\theta(s_0) = \theta_0$.

(ii) For $H \neq 0$ (resp. $H = 0$) there are global integral curves of M_1 (resp. A_1).

Proof. (i) With the help of (VIII.6.), we can see that the functions in (VIII.8.) are continuous and bounded everywhere. Therefore, the existence theorem of Peano (cf. Kamke [8], page 126) holds.

(ii) In (VII.4.), $\boldsymbol{q}$ and $\boldsymbol{p}$ were expressed as functions of $\boldsymbol{M}$, $\boldsymbol{L}$, and θ. In (VIII.6.), $\boldsymbol{M}$ and $\boldsymbol{L}$ were expressed as functions of s. Because of (VIII.8.) and (VIII.9.) (i), the desired function $\theta(s)$ exists. With that, we have $\boldsymbol{q}$ and $\boldsymbol{p}$ as functions of the integral curve parameter s. ■

IX. Stereographic Projection and global SO(4) (resp. SO(3,1)) Symmetry

In the case of quantum mechanics, by using a stereographic projection in momentum space ($\boldsymbol{p}$ space), it is possible to construct a representation of the group SO(4) as a symmetry group in the Hilbert space of the hydrogen atom (cf. Bander and Itzykson [3]). This method goes back to Fock. In his discussion of compactification of the phase space, Moser [10] provides a classical mechanical analog of this projection. He maps the hyper surface of constant energy of phase space onto a sub-bundle of the tangent bundle of the sphere S^3. When restricted to momentum space, this mapping is also a stereographic projection. The image space thus has a special symmetry property. We make use of that, for $H < 0$ (resp. for $H > 0$), in order to map the energy surface onto a sub-bundle of the tangent bundle of a sphere (resp. of a hyperboloid). This makes it possible to solve the differential equations for the integral curves of M_1 for $H \neq 0$. We didn't find the corresponding solution for $H = 0$.

(IX.1.) Definition.

(i) For $H = H_0 < 0$:

$$\xi_0 = \frac{\|\boldsymbol{p}\|^2 + 2H_0\mu}{\|\boldsymbol{p}\|^2 - 2H_0\mu},$$

$$\xi_k = \frac{-2\sqrt{-2H_0\mu}\,p_k}{\|\boldsymbol{p}\|^2 - 2H_0\mu},$$

$$\eta_0 = \frac{-\sqrt{-2H_0}}{\sqrt{\mu}}(\boldsymbol{q} \cdot \boldsymbol{p}),$$

$$\eta_k = \left(\frac{\|\boldsymbol{p}\|^2 - 2H_0\mu}{2\mu}\right) q_k - \frac{(\boldsymbol{q} \cdot \boldsymbol{p})}{\mu} p_k,$$

(ii) For $H = H_0 > 0$:

$$\xi_0 = \frac{\|\boldsymbol{p}\|^2 + 2H_0\mu}{\|\boldsymbol{p}\|^2 - 2H_0\mu},$$

$$\xi_k = \frac{-2\sqrt{2H_0\mu}\,p_k}{\|\boldsymbol{p}\|^2 - 2H_0\mu},$$

$$\eta_0 = \frac{-\sqrt{2H_0}}{\sqrt{\mu}}(\boldsymbol{q} \cdot \boldsymbol{p}),$$

$$\eta_k = -\left(\frac{\|\boldsymbol{p}\|^2 - 2H_0\mu}{2\mu}\right) q_k + \frac{(\boldsymbol{q} \cdot \boldsymbol{p})}{\mu} p_k.$$

Remark. For $H = H_0 > 0$, we have $\|\boldsymbol{p}\|^2 - 2H_0\mu = 2\mu/\|\boldsymbol{q}\| > 0$. As a result, $\xi_0 \geq 1$ and the transformation is on the upper piece of the hyperboloid.

(IX.2.) Corollary.

(i) For $H = H_0 < 0$:

$$p_k = -\sqrt{-2H_0\mu}\left(\frac{\xi_k}{1-\xi_0}\right),$$

$$q_k = (1/{-2H_0})(\eta_k(1-\xi_0) + \xi_k\eta_0),$$

(ii) For $H = H_0 > 0$:

$$p_k = +\sqrt{2H_0\mu}\left(\frac{\xi_k}{1-\xi_0}\right),$$

$$q_k = (1/2H_0)(\eta_k(1-\xi_0) + \xi_k\eta_0).$$

(IX.3.) Corollary.

(i) For $H = H_0 < 0$:

$\|\boldsymbol{p}\|^2 = (-2H_0\mu)\left(\frac{1+\xi_0}{1-\xi_0}\right)$,

$\frac{1}{\|\boldsymbol{q}\|} = \frac{-2H_0}{1-\xi_0}$,

$A_1 = \xi_1\eta_0 - \xi_0\eta_1$,

${\xi_0}^2 + \sum_{k=1}^{3} {\xi_k}^2 = 1$,

$\xi_0\eta_0 + \sum_{k=1}^{3} \xi_k\eta_k = 0$,

${\eta_0}^2 + \sum_{k=1}^{3} {\eta_k}^2 = 1$,

(ii) For $H = H_0 > 0$*:*

$\|\boldsymbol{p}\|^2 = (-2H_0\mu)\left(\frac{1+\xi_0}{1-\xi_0}\right)$,

$\frac{1}{\|\boldsymbol{q}\|} = \frac{-2H_0}{1-\xi_0}$,

$A_1 = \xi_0\eta_1 - \xi_1\eta_0$,

${\xi_0}^2 - \sum_{k=1}^{3} {\xi_k}^2 = 1$,

$\xi_0\eta_0 - \sum_{k=1}^{3} \xi_k\eta_k = 0$,

${\eta_0}^2 - \sum_{k=1}^{3} {\eta_k}^2 = -1$.

Remark. For $\mathrm{H} < 0$, the ξs form a 3-dimensional sphere $S^3 \subset \mathbb{R}^4$, and for $\mathrm{H} > 0$ a hyperboloid. In both cases, the ηs are tangential to this hypersurface, and fulfill one additional condition. As a result, the energy surface is mapped onto a 5-dimensional subspace of $\mathbb{R}^8$ that has a canonical action of the group SO(4) resp. SO(3,1). The fact that this is a tangent space could be used to transport the symplectic structure. However, we won't pursue this possibility, because it appears to be very complicated. Instead, we will translate the differential equations for an integral curve of M_1 directly.

(IX.4.) Theorem. *Let* $c: \mathbb{R} \to M | s \mapsto (\boldsymbol{q}, \boldsymbol{p})$ *be an integral curve of* M_1 *and* ξ_i, η_i *defined as in (IX.1.). Then:*

(i) For $H < 0$*:*

$\xi_0' = +\xi_1 - \eta_0\left(\frac{\eta_1}{1-\xi_0}\right) = +\xi_1 - \eta_0{\eta_0}'$,

$\xi_1' = -\xi_0 - \eta_1\left(\frac{\eta_1}{1-\xi_0}\right) = -\xi_0 - \eta_1{\eta_0}'$,

$\xi_2' = -\eta_2\left(\frac{\eta_1}{1-\xi_0}\right) = -\eta_2{\eta_0}'$,

$\xi_3' = -\eta_3\left(\frac{\eta_1}{1-\xi_0}\right) = -\eta_3{\eta_0}'$,

$\eta_0' = +\eta_1 + \xi_0\left(\frac{\eta_1}{1-\xi_0}\right) = \frac{\eta_1}{1-\xi_0} = +\eta_1 + \xi_0{\eta_0}'$,

$\eta_1' = -\eta_0 + \xi_1\left(\frac{\eta_1}{1-\xi_0}\right) = -\eta_0 + \xi_1{\eta_0}'$,

$\eta_2' = +\xi_2\left(\frac{\eta_1}{1-\xi_0}\right) = +\xi_2{\eta_0}'$,

$\eta_3' = +\xi_3\left(\frac{\eta_1}{1-\xi_0}\right) = +\xi_3{\eta_0}'$.

(ii) For $H > 0$*:*

$\xi_0' = -\xi_1 - \eta_0\left(\frac{\eta_1}{1-\xi_0}\right) = -\xi_1 + \eta_0{\eta_0}'$,

$\xi_1' = -\xi_0 - \eta_1\left(\frac{\eta_1}{1-\xi_0}\right) = -\xi_0 + \eta_1\eta_0{}'$,

$\xi_2' = -\eta_2\left(\frac{\eta_1}{1-\xi_0}\right) = +\eta_2\eta_0{}'$,

$\xi_3' = -\eta_3\left(\frac{\eta_1}{1-\xi_0}\right) = +\eta_3\eta_0{}'$,

$\eta_0' = -\eta_1 - \xi_0\left(\frac{\eta_1}{1-\xi_0}\right) = \frac{-\eta_1}{1-\xi_0} = -\eta_1 + \xi_0\eta_0{}'$,

$\eta_1' = -\eta_0 - \xi_1\left(\frac{\eta_1}{1-\xi_0}\right) = -\eta_0 + \xi_1\eta_0{}'$,

$\eta_2' = -\xi_2\left(\frac{\eta_1}{1-\xi_0}\right) = +\xi_2\eta_0{}'$,

$\eta_3' = -\xi_3\left(\frac{\eta_1}{1-\xi_0}\right) = +\xi_3\eta_0{}'$.

Proof. In each case, the proof consists of a direct application of the definitions and a lengthy calculation. We will sketch the proof for ξ_0' and $H < 0$.

$$\xi_0' = \frac{2(\boldsymbol{p}\cdot\boldsymbol{p}')}{\|\boldsymbol{p}\|^2 - 2H_0\mu} - \frac{2(\boldsymbol{p}\cdot\boldsymbol{p}')(\|\boldsymbol{p}\|^2 + 2H_0\mu)}{(\|\boldsymbol{p}\|^2 - 2H_0\mu)^2} = \frac{2(\boldsymbol{p}\cdot\boldsymbol{p}')(-4H_0\mu)}{(\|\boldsymbol{p}\|^2 - 2H_0\mu)^2}.$$

Because of (VIII.1.), we have:

$$(\boldsymbol{p}\cdot\boldsymbol{p}') = \sqrt{\frac{\mu}{2}}(-H_0)^{-3/2}\left[\frac{A_1}{2\|\boldsymbol{q}\|^3}(-\boldsymbol{q}\cdot\boldsymbol{p}) + H_0\left(\frac{p_1}{\|\boldsymbol{q}\|} - \frac{q_1(\boldsymbol{q}\cdot\boldsymbol{p})}{\|\boldsymbol{q}\|^3}\right)\right].$$

Now we apply (IX.2.) and (IX.3.), and get:

$(\boldsymbol{p}\cdot\boldsymbol{p}') = -2H_0\mu(1/(1-\xi_0)^3)(\xi_1(1-\xi_0) - \eta_0\eta_1)$, and

$\xi_0' = (1/(1-\xi_0))(\xi_1 - \xi_0\xi_1 - \eta_0\eta_1)$. ■

(IX.5.) Theorem. *Let* $c\colon \mathbb{R} \to M|s \mapsto (\boldsymbol{q},\boldsymbol{p})$ *be an integral curve of* M_1 *and* ξ_i, η_i *defined as in (IX.1.). Then:*

(i) For $H < 0$*:*

$\xi_0 + i\xi_1 = Be^{-is}cos(\eta_0 + \alpha)$,

$\xi_2 = B_2 cos(\eta_0 + \alpha_2)$,

$\xi_3 = B_3 cos(\eta_0 + \alpha_3)$,

$\eta_0 + i\eta_1 = Be^{-is}sin(\eta_0 + \alpha)$,

$\eta_2 = B_2 sin(\eta_0 + \alpha_2)$,

$\eta_3 = B_3 sin(\eta_0 + \alpha_3)$,

$\eta_0 = (1/2)\left[Be^{-is}\,sin(\eta_0 + \alpha) + \bar{B}e^{+is}\,sin(\eta_0 + \bar{\alpha})\right]$,

$B, \alpha \in \mathbb{C}$ *and* $B_2, B_3, \alpha_2, \alpha_3 \in \mathbb{R}$ *are the integration constants, with*

(1) $1 = \|B\|^2 sin(\eta_0 + \alpha)sin(\eta_0 + \bar{\alpha}) + {B_2}^2 sin^2(\eta_0 + \alpha_2) + {B_3}^2 sin^2(\eta_0 + \alpha_3)$,

(2) $1 = \|B\|^2 cos(\eta_0 + \alpha)cos(\eta_0 + \bar{\alpha}) + {B_2}^2 cos^2(\eta_0 + \alpha_2) + {B_3}^2 cos^2(\eta_0 + \alpha_3)$,

$\Rightarrow 2 = \|B\|^2\,cosh(2Im\alpha) + B_2^2 + B_3^2$,

(3) $0 = \|B\|^4 + B_2^4 + B_3^4 + 2\|B\|^2 B_2^2\,cos(2Re\alpha - 2\alpha_2) + 2\|B\|^2 B_3^2\,cos(2Re\alpha - 2\alpha_3)$
$+2B_2^2 B_3^2\,cos(2\alpha_2 - 2\alpha_3)$,

$A_1 = (-1/2)\|B\|^2\,sinh(2Im\alpha)$.

(ii) For $H > 0$*:*

$\xi_0 + \xi_1 = B_0 e^{-s} cosh(\eta_0 + \alpha_0)$,

$\xi_0 - \xi_1 = B_1 e^{+s} cosh(\eta_0 + \alpha_1)$,

$\xi_2 = B_2 cosh(\eta_0 + \alpha_2)$,

$\xi_3 = B_3 cosh(\eta_0 + \alpha_3)$,

$\eta_0 + \eta_1 = B_0 e^{-s} sinh(\eta_0 + \alpha_0)$,

$\eta_0 - \eta_1 = B_1 e^{+s} sinh(\eta_0 + \alpha_1)$,

$\eta_2 = B_2 sinh(\eta_0 + \alpha_2)$,

$\eta_3 = B_3 sinh(\eta_0 + \alpha_3)$,

$\eta_0 = (1/2)[B_0 e^{-s}\, sinh(\eta_0 + \alpha_0) + B_1 e^{+s}\, sinh(\eta_0 + \alpha_1)]$,

$B_0, B_1, B_2, B_3, \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ *are the integration constants, with*

(1) $1 = B_0 B_1 cosh(\eta_0 + \alpha_0) cosh(\eta_0 + \alpha_1) - {B_2}^2 cosh^2(\eta_0 + \alpha_2) - {B_3}^2 cosh^2(\eta_0 + \alpha_3)$,

(2) $1 = -B_0 B_1 sinh(\eta_0 + \alpha_0) sinh(\eta_0 + \alpha_1) + {B_2}^2 sinh^2(\eta_0 + \alpha_2) + {B_3}^2 sinh^2(\eta_0 + \alpha_3)$,

$\Rightarrow 2 = B_0 B_1 cosh(\alpha_0 - \alpha_1) - B_2^2 - B_3^2$,

(3) $0 = B_0^2 B_1^2 - B_2^4 - B_3^4 - 2B_0 B_1 B_2^2\, cosh(\alpha_0 + \alpha_1 - 2\alpha_2)$

$\quad -2B_0 B_1 B_3^2\, cosh(\alpha_0 + \alpha_1 - 2\alpha_3) + B_2^2 B_3^2 cosh(2\alpha_2 - 2\alpha_3)$,

$A_1 = (1/2) B_0 B_1\, sinh(\alpha_0 - \alpha_1)$.

Note. There is a typographical error in the original version. For $\mathrm{H} < 0$ and $\mathrm{H} > 0$, in the expressions (1) and (2), the η_0 was missing. We have corrected it here (more details in the supplementary information).

Proof. The formulas for ξ_i, η_i are the solutions of the differential equations given in (IX.4.), and the additional conditions on the integration constants follow from the additional conditions in (IX.3.). ■

Remark. Because of $\xi_0 \geq 1$ and $\xi_0^2 - \xi_1^2 \geq 1$, we have $\xi_0 + \xi_1 \geq 0$. Then (IX.5.) (ii) leads to $B_0 \geq 0$ and $B_1 \geq 0$.

(IX.6.) Theorem. *For* $H < 0$ *and* $H > 0$*, the implicit equation* $\eta_0 = \eta_0(s)$ *in (IX.5.) has exactly one solution.*

Proof. For $\mathrm{H} < 0$, the equation is:

$\eta_0 = (1/2)\left[B e^{-is} \sin(\eta_0 + \alpha) + \bar{B} e^{+is} \sin(\eta_0 + \bar{\alpha})\right]$. For fixed *s*, we will search for the intersection of the two functions that are on the left and the right of the equality sign. The left side is the straight line through zero with a slope of $+1$, and the right side is a periodic, bounded function of η_0. Therefore, there is at least one intersection. The uniqueness stems from the fact that the slope of the function on the right side is always ≤ 1. If we denote this slope with *S*, the we have:

$S = (1/2)\left[B e^{-is}\, cos(\eta_0 + \alpha) + \bar{B} e^{+is}\, cos(\eta_0 + \bar{\alpha})\right] \Rightarrow$

$$S^2 = (1/4)\begin{bmatrix} \|B\|^2 cos(\eta_0 + \alpha) cos(\eta_0 + \bar{\alpha}) + \|B\|^2 cos(\eta_0 + \bar{\alpha}) cos(\eta_0 + \alpha) \\ +B^2 e^{-2is} cos^2(\eta_0 + \alpha) + \bar{B}^2 e^{+2is} cos^2(\eta_0 + \bar{\alpha}) \end{bmatrix},$$

$S^2 = (1/4)[2(\xi_0 + i\xi_1)(\xi_0 - \xi_1) + (\xi_0 + i\xi_1)^2 + (\xi_0 - i\xi_1)^2] = \xi_0^2 \leq 1$.

(ii) For $\mathrm{H} > 0$, the equation is:

$\eta_0 = (1/2)[B_0 e^{-s}\sinh(\eta_0+\alpha_0) + B_1 e^{+s}\sinh(\eta_0+\alpha_1)]$. This time, the right side is a function that behaves like $e^{+\eta_0}$ (resp. $-e^{-\eta_0}$) for $\eta_0 \gg 0$ (resp. $\eta_0 \ll 0$), since $B_0 \geq 0$ and $B_1 \geq 0$. As a result, there is at least one intersection. The uniqueness stems from the fact that the slope of the function on the right side is always ≥ 1. If we denote this slope with S, the we have:

$$S = (1/2)[B_0 e^{-s}\cosh(\eta_0+\alpha_0) + B_1 e^{+s}\cosh(\eta_0+\alpha_1)].$$

$$S^2 = (1/4)\begin{bmatrix} B_0^2 e^{-2s}\cosh^2(\eta_0+\alpha_0) + B_1^2 e^{+2s}\cosh^2(\eta_0+\alpha_0) \\ +2B_0B_1\cosh(\eta_0+\alpha_0)\cosh(\eta_0+\alpha_1) \end{bmatrix},$$

$$S^2 = (1/4)[(\xi_0+\xi_1)^2 + (\xi_0-\xi_1)^2 + 2(\xi_0+\xi_1)(\xi_0-\xi_1)] = \xi_0^2 \geq 1.$$

In both cases, the slope can only have a value of 1 at isolated points, since the function is analytical. ■

(IX.7.) Definition.

For $H < 0$:

$$\psi := \sqrt{\frac{-2H}{\mu}}(\boldsymbol{q}\cdot\boldsymbol{p}),$$

$$x_0 := \left(\frac{\|\boldsymbol{q}\|\cdot\|\boldsymbol{p}\|^2}{\mu} - 1\right)\cos\psi + \sqrt{\frac{-2H}{\mu}}(\boldsymbol{q}\cdot\boldsymbol{p})\sin\psi,$$

$$x_k := \sqrt{\frac{-2H}{\mu}}\|\boldsymbol{q}\|p_k\cos\psi + \left(\frac{q_k}{\|\boldsymbol{q}\|} - \frac{(\boldsymbol{q}\cdot\boldsymbol{p})p_k}{\mu}\right)\sin\psi,$$

$$y_0 := (\boldsymbol{q}\cdot\boldsymbol{p})\cos\psi - \sqrt{\frac{\mu}{-2H}}\left(\frac{\|\boldsymbol{q}\|\cdot\|\boldsymbol{p}\|^2}{\mu} - 1\right)\sin\psi,$$

$$y_k := \sqrt{\frac{\mu}{-2H}}\left(\frac{q_k}{\|\boldsymbol{q}\|} - \frac{(\boldsymbol{q}\cdot\boldsymbol{p})p_k}{\mu}\right)\cos\psi - (\|\boldsymbol{q}\|p_k)\sin\psi.$$

For $H > 0$:

$$\psi := \sqrt{\frac{2H}{\mu}}(\boldsymbol{q}\cdot\boldsymbol{p}),$$

$$x_0 := \left(\frac{\|\boldsymbol{q}\|\cdot\|\boldsymbol{p}\|^2}{\mu} - 1\right)\cosh\psi - \sqrt{\frac{2H}{\mu}}(\boldsymbol{q}\cdot\boldsymbol{p})\sinh\psi,$$

$$x_k := \sqrt{\frac{2H}{\mu}}\|\boldsymbol{q}\|p_k\cosh\psi + \left(\frac{q_k}{\|\boldsymbol{q}\|} - \frac{(\boldsymbol{q}\cdot\boldsymbol{p})p_k}{\mu}\right)\sinh\psi,$$

$$y_0 := (\boldsymbol{q}\cdot\boldsymbol{p})\cosh\psi - \sqrt{\frac{\mu}{2H}}\left(\frac{\|\boldsymbol{q}\|\cdot\|\boldsymbol{p}\|^2}{\mu} - 1\right)\sinh\psi,$$

$$y_k := -\sqrt{\frac{\mu}{2H}}\left(\frac{q_k}{\|\boldsymbol{q}\|} - \frac{(\boldsymbol{q}\cdot\boldsymbol{p})p_k}{\mu}\right)\cosh\psi - (\|\boldsymbol{q}\|p_k)\sinh\psi.$$

Note. We have included the case of negative energy, using the definition from a later publication (Ligon and Schaaf, 1976).

(IX.8.) Corollary.

For $H < 0$:

$$\textstyle\sum_{i=0}^{3} x_i^2 = 1,$$

$$\textstyle\sum_{i=0}^{3} x_i y_i = 0,$$

$$\textstyle\sum_{i=0}^{3} y_i^2 = -\frac{\mu}{2H}.$$

For $H > 0$:

$x_0^2 - \sum_{i=1}^3 x_i^2 = 1,$

$x_0 y_0 - \sum_{i=1}^3 x_i y_i = 0,$

$y_0^2 - \sum_{i=1}^3 y_i^2 = -\frac{\mu}{2H}.$

Proof. This is a long but elementary calculation. ■

(IX.9.) Theorem.

(i) For $H < 0$, along an integral curve of M_1:

$x_0' = -x_1, \qquad y_0' = -y_1,$

$x_1' = +x_0, \qquad y_1' = +y_0,$

$x_2' = 0, \qquad y_2' = 0,$

$x_3' = 0, \qquad y_3' = 0.$

(ii) For $H < 0$, along an integral curve of L_1:

$x_0' = 0, \qquad y_0' = 0,$

$x_1' = 0, \qquad y_1' = 0,$

$x_2' = -x_3, \qquad y_2' = -y_3,$

$x_3' = +x_2, \qquad y_3' = +y_2.$

(i) For $H > 0$, along an integral curve of M_1:

$x_0' = x_1, \qquad y_0' = y_1,$

$x_1' = +x_0, \qquad y_1' = +y_0,$

$x_2' = 0, \qquad y_2' = 0,$

$x_3' = 0, \qquad y_3' = 0.$

(ii) For $H > 0$, along an integral curve of L_1:

$x_0' = 0, \qquad y_0' = 0,$

$x_1' = 0, \qquad y_1' = 0,$

$x_2' = -x_3, \qquad y_2' = -y_3,$

$x_3' = +x_2, \qquad y_3' = +y_2.$

Proof. The proof is analogous to (IX.4.), and just as long.

Remark. The analogous equations for $M_2, M_3, L_2,$ and L_3 are obtained via cyclic permutation of the indices. Then, this set of equations is immediately integrable, and provide the canonical action of SO(4).

X. Concluding Remarks

The first substantial part of this investigation of the Kepler problem was the search for the integral curves of the Hamiltonian function H. This was achieved most elegantly by utilizing the "maximal" symmetry of the problem. With this symmetry, it is possible to find the trajectories of H via a very simple method. The symmetry is related to the conservation of angular momentum and of the Runge-Lenz vector, which form a Lie algebra with the Poisson bracket.

The second part was a closer investigation of the symmetry. We searched for the transformation that is related to the conservation of the Runge-Lenz vector, i.e. we searched for the integral curves of $\boldsymbol{M}$. It was not possible to solve this problem the same way as the first one, but the utilization of the symmetry was essential. The energy surface was mapped to a space that has an especially simple action of the desired group. This "simplest", or "canonical" action of $\mathrm{SO}(4)$ on $T_1(S^3)$ was, however, **not** the action of $\mathrm{SO}(4)$ that has the functions $\boldsymbol{L}$ and $\boldsymbol{M}$ as its Lie algebra. However, the differential equations for the integral curves of $\boldsymbol{M}$ that are transferred by this transformation were easy to solve. One deviation from the "canonical" action of $\mathrm{SO}(4)$ is the "frequency" η_0, which is defined by a transcendental equation.

The quantum-mechanical case (hydrogen atom) is in many ways analogous to the classical mechanical problem, and has been discussed, e.g. by Bander and Itzykson [3]. The Runge-Lenz vector is defined as a hermitic operator and provides, with its commutation relations, a representation of the same Lie algebra as in the classical case. This procedure, which goes back to Fock, defines a stereographic projection in momentum space. The resulting Schrödinger equation is then, for $H < 0$, invariant under $\mathrm{SO}(4)$ in an obvious way. The sphere S^3 that appears in the stereographic projection as image space has a canonical action of the group $\mathrm{SO}(4)$ (rotation of the sphere), which generates a unitary representation of the group $\mathrm{SO}(4)$. It has also been shown that the self-adjoint operators that serve as infinitesimal generators of this representation are precisely the operators of the angular momentum and the Runge-Lenz vector. The existence of such a representation of the group is closely related to the existence of a global action of the group in the classical case (cf. Kostant [9;(2.10.1)]. That is an application of "Kostantification", a functor from the category of Hamiltonian systems with Poisson bracket Lie algebra in the category of Hilbert spaces and linear operators with Lie algebra of commutation relations.

Literature

1 Abraham, R. H. and J. E. Marsden (1967). "Foundations of Mechanics: a Mathematical Exposition of Classical Mechanics With an Introduction to the Qualitative Theory of Dynamical Systems and Applications to the Three-Body Problem." Benjamin, New York.

2 Bacry, H. (1966). "The de Sitter group $L_{4,1}$ and the Bound States of Hydrogen Atom." Il Nuovo Cimento A (1965-1970) 41(2): 222-234.

3 Bander, M. and C. Itzykson (1966). "Group Theory and the Hydrogen Atom (I&II)." Reviews of Modern Physics 38(2): 330-345&346-358.

4 Dieudonné, J. (1971). "Elements d'Analyse."

5 Dirac, P. (1947). "The Principles of Quantum Mechanics." The International Series of Monographs on Physics, Oxford: Clarendon Press.

6 Györgyi, G. (1968). "Kepler's Equation, Fock Variables, Bacry's Generators and Dirac Brackets." Il Nuovo Cimento A Series 10 53(3): 717-736 and 62A: 449.

7 Jost, R. (1964). "Poisson Brackets (an Unpedagogical Lecture)." Reviews of Modern Physics 36(2): 572-579.

8 Kamke, E. (1956). Differentialgleichungen Reeller Funktionen, 3. Auflage; Akademische Verlagsgesellchaft, Leipzig.

9 Kostant, B. (1970). Quantization and Unitary Representations. Lectures in Modern Analysis and Applications III. C. T. Taam. Berlin, Heidelberg, Springer Berlin Heidelberg: 87-208.

10 Moser, J. (1970). "Regularization of Kepler's Problem and the Averaging Method on a Manifold." Communications on Pure and Applied Mathematics 23(4): 609-636.

11 Pauli, W. (1953). "On the Hamiltonian Structure of Non-Local Field Theories." Il Nuovo Cimento (1943-1954) 10(5): 648-667.

12 Rottmann, K. (1960). "Mathematische Formelsammlung." Bibliographisches Institut-Hochschultaschenbuecher, Mannheim: Bibliographisches Institut, 2. Auflage.

13 Ryshik, I. u. IS Gradstein (1957). Summen-, Produkt-und Integraltafeln. Berlin: VEB, Deutscher Verlag der Wissenschaften.

14 Segal, I. (1960). "Quantization of Nonlinear Systems." Journal of Mathematical Physics 1(6): 468-488.