CWI Syllabi

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

CWI Syllabus 13

Proceedings Seminar 1983-1985 Mathematical structures in field theories, Vol.2

M.J. Bergvelt
G.M. Tuynman
A.P.E. ten Kroode

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification: 70HXX, 53BXX.
ISBN 90 6196 317 6
NUGI-code: 811

Printed in the Netherlands

Table of Contents

Preface

These proceedings cover part of the lectures given in the seminar 'Mathematical Structures in Field Theories', held at the University of Amsterdam during the academic years 1983-1984 and 1984-1985 (see CWI-Syllabi 2, 6 and 8).

Chapter 1 by G.M. Tuynman gives an introduction to classical mechanics and symplectic geometry and is an introduction to the next two chapters.

M.G. Bergvelt treats in the second chapter Yang-Mills theory as a classical albeit singular dynamical system; the mathematical frame work is in terms of differential geometry and the paper is an application of the work by Gotay, Nester and Hinds to the Yang-Mills system.

The third chapter by A.P.E. ten Kroode is devoted to the geometrical description of the Toda lattice. This lattice is described as a Hamiltonian system on a co-adjoint orbit in the dual of a Lie algebra. The symplectic structure is the Kostant-Kirillov symplectic form.

We thank the Centre for Mathematics and Computer Science again for the excellent technical production of these proceedings.

The editors
E.M. de Jager
H.G.J. Pijls

An Introduction to Classical Mechanics

and

Symplectic Geometry

G.M. Tuynman

Mathematical Institute, University of Amsterdam

§1. NEWTONIAN MECHANICS

The equations of motion of a system of N particles are given by Newton's second law:

$$(1) \qquad m_i \frac{d^2q^i}{dt^2}(t) = F_i(q(t), \frac{dq}{dt}(t)) \qquad i = 1,\ldots,3N$$

where we have introduced the $n = 3N$ coordinates $q^1, \ldots, q^n$ on $\mathbb{R}^{3N}$ which denote the positions of the N particles in space; m_i denotes the mass of a particle and F_i denotes the force exerted on the "i-th" particle.

In general F_i can depend on the positions and the velocities of the particles and explicitly on the time t, but in these notes we restrict ourselves to the so-called time independent systems in which F_i does not depend on t explicitly.

<u>EXAMPLE 1</u>: a particle in a constant "vertical" gravity field:

$$(F_i) = (0,0,-mg) \qquad (i=1,2,3) \qquad m_i = m$$

<u>EXAMPLE 2</u>: a particle in a gravity well:

$$(F_i) = -m\gamma r^{-3} \cdot (q^1,q^2,q^3), \qquad r^2 = (q^1)^2 + (q^2)^2 + (q^3)^2.$$

<u>EXAMPLE 3</u>: a charged particle in a (static) electromagnetic field with the usual Lorentz force:

$$\vec{F}(\vec{q}, \frac{d\vec{q}}{dt}) = e\{\vec{E}(\vec{q}) + \frac{d\vec{q}}{dt} \wedge \vec{B}(\vec{q})\}$$

where e is the electric charge of the particle, $\vec{E}$ the electric field, $\vec{B}$ the magnetic field and "$\wedge$" the vector product in $\mathbb{R}^3$.

<u>EXAMPLE 4</u>: a damped oscillator (in 1-dimension):

$$F(q,\frac{dq}{dt}) = -kq - \alpha\frac{dq}{dt}.$$

§2. THE LAGRANGE FORMALISM

The Lagrange formulation of classical mechanics starts with the assumption that the force F_i can be derived from a potential function $V(q,\dot{q})$ by

$$(2) \qquad F_i\left(q(t),\frac{dq}{dt}(t)\right) = -\frac{\partial V}{\partial q^i}\left(q(t),\frac{dq}{dt}(t)\right) + \frac{d}{dt}\frac{\partial V}{\partial \dot{q}^i}\left(q(t),\frac{dq}{dt}(t)\right)$$

where V is a function of $2n$ variables q^i, $\dot{q}^i$ (it should be noted that most systems which occur in practice, satisfy this assumption). Now Newton's equations become:

$$(3) \qquad m_i\frac{d^2q^i}{dt^2}(t) = \frac{d}{dt}\frac{\partial V}{\partial \dot{q}^i}\left(q(t),\frac{dq}{dt}(t)\right) - \frac{\partial V}{\partial q^i}\left(q(t),\frac{dq}{dt}(t)\right).$$

EXAMPLE 1: $V(q,\dot{q}) = mg\,q^3$

EXAMPLE 2: $V(q,\dot{q}) = -m\gamma r^{-1}$

EXAMPLE 3: $V(q,\dot{q}) = e\{\phi(q) - \dot{q}^i A_i(q)\}$
where A_i and ϕ are defined by the equations $\vec{E} = -\mathrm{grad}\,\phi$, $\vec{B} = \mathrm{rot}\,\vec{A}$.

EXAMPLE 4: for this system no such potential function exists.

It is well-known that the equations (3) are not invariant under changes of coordinates, e.g. in example 2 the equation

$$m\frac{d^2r}{dt^2} = -\frac{\partial V}{\partial r}$$

is not compatible with the (Newtonian) equations of motion (3).
In order to derive a form of the equations of motion which is invariant under changes of coordinates we introduce the kinetic energy function T by

$$T(q,\dot{q}) = \sum_{i=1}^{n} \tfrac{1}{2}m_i(\dot{q}^i)^2$$

and we define the Lagrange function (or Lagrangian) L by

$$L(q,\dot{q}) = T(q,\dot{q}) - V(q,\dot{q}).$$

With these definitions we can rewrite (3) as:

(4a) $$\frac{\partial L}{\partial q^i}(q(t),\dot{q}(t)) = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}(q(t),\dot{q}(t))$$

(4b) $$\dot{q}^i(t) = \frac{dq^i}{dt}(t)$$

which are called the Euler-Lagrange equations of motion.

A few remarks can be made at this point. Firstly, equation (4b) strongly suggests that the space $\mathbb{R}^{2n}$ is the tangent space $T\mathbb{R}^n$ where q are coordinates on $\mathbb{R}^n$ and $\dot{q}$ the associated coordinates in the fibres (tangent space):

$(q,\dot{q})$ represents the tangent vector $\dot{q}^i \frac{\partial}{\partial q^i}$ at $\dot{q}$.

Secondly, with the above indentification of $\mathbb{R}^{2n}$ with $T\mathbb{R}^n$ we see that T, V and L are functions on $T\mathbb{R}^n$ and that the kinetic energy defines a metric on $\mathbb{R}^n$.

The usefulness of the above definitions is expressed by the following proposition.

PROPOSITION 2.1: *the form of the Euler-Lagrange equations is invariant under changes of coordinates; more precisely if* $(\hat{q})$ *is a different set of coordinates on* $\mathbb{R}^n$ *with associated coordinates* $\dot{\hat{q}}$ *in* $T\mathbb{R}^n$ *then* (4) *is equivalent with*

(5) $$\begin{cases} \dfrac{\partial L}{\partial \hat{q}^i}(\hat{q}(t),\dot{\hat{q}}(t)) = \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{\hat{q}}^i}(\hat{q}(t),\dot{\hat{q}}(t) \\ \dot{\hat{q}}^i(t) = \dfrac{d\hat{q}^i}{dt}(t) \end{cases}$$

PROOF: we have coordinates $\hat{q}^i = \hat{q}^i(q)$ hence the associated coordinates in the fibres of $T\mathbb{R}^n$ transform according to

$$(6) \qquad \dot{\hat{q}}^i = \frac{\partial \hat{q}^i}{\partial q^j}\dot{q}^j$$

Now $$\frac{\partial \hat{q}^i}{\partial q^j}(q)\,\dot{q}^j = \dot{\hat{q}}^i \overset{(5)}{=} \frac{d\hat{q}^i}{dt} = \frac{\partial \hat{q}^i}{\partial q^j}\frac{dq^j}{dt}$$

hence $\dot{q}^j = \frac{dq^j}{dt}$ because $\frac{\partial \hat{q}^i}{\partial q^j}$ is invertible.

Furthermore, we have the 3 equalities:

$$\frac{\partial L}{\partial q^i} = \frac{\partial L}{\partial \hat{q}^j}\frac{\partial \hat{q}^j}{\partial q^i} + \frac{\partial L}{\partial \dot{\hat{q}}^j}\frac{\partial \dot{\hat{q}}^j}{\partial q^i}$$

$$\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial \hat{q}^j}\frac{\partial \hat{q}^j}{\partial \dot{q}^i} + \frac{\partial L}{\partial \dot{\hat{q}}^j}\frac{\partial \dot{\hat{q}}^j}{\partial \dot{q}^i} \overset{(6)}{=} \frac{\partial L}{\partial \dot{\hat{q}}^j}\frac{\partial \hat{q}^j}{\partial q^i}$$

$$\frac{d}{dt}\frac{\partial \hat{q}^j}{\partial q^i} = \frac{\partial^2 \hat{q}^j}{\partial q^k \partial q^i}\frac{dq^k}{dt} = \frac{\partial^2 \hat{q}^j}{\partial q^i \partial q^k}\dot{q}^k = \frac{\partial}{\partial q^i}\left(\frac{\partial \hat{q}^j}{\partial q^k}\dot{q}^k\right) = \frac{\partial \dot{\hat{q}}^j}{\partial q^i}\,.$$

Combining these 3 equalities we get $\frac{\partial L}{\partial q^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}$. This proves that (5) implies (4); the other implication is proved analogous. QED

Till this point we supposed that the "position" of our system could be any point of $\mathbb{R}^n$ (the so-called configuration space). More generally one considers systems in which the position of the physical system is a point of an arbitrary manifold Q for which the motion is described by a Lagrangian $L: TQ \to \mathbb{R}$. Such systems occur when there exist constraints which restrict the position of the system (in $\mathbb{R}^n$) to a submanifold Q of $\mathbb{R}^n$. Then the position of the system is described by a point of Q and the velocities $\dot{q}(t) = \frac{dq}{dt}$ will be tangent to Q. One can prove (e.g. see [Arnold] and references cited there) that the equations of motion remain the same:

$$\frac{\partial L}{\partial \hat{q}^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{\hat{q}}^i}\,, \qquad \frac{d\hat{q}^i}{dt} = \dot{\hat{q}}^i$$

but now $\hat{q}$ are (local) coordinates on Q with associated coordinates $\dot{\hat{q}}$ in TQ.

REMARK: proposition 2.1 guarantees us that the time evolution of the "position" of the physical system ("position" seen as a point of TQ) as described by the equations (5) is independent of the chosen coordinate system on Q hence the time-evolution is a well-defined process.

REMARK: on $\mathbb{R}^n$ the kinetic energy T defined a metric; if Q is a submanifold of $\mathbb{R}^n$ to which a mechanical system is restricted, then both T and V are functions on TQ and T (again) defines a metric on Q (which is the metric induced by the metric on $\mathbb{R}^n$ defined by T).

EXAMPLE: two particles of mass m_1, m_2 joined by a rod of length 1 then $\mathbb{R}^6 \supset Q \cong \mathbb{R}^3 \times S^2$. On $\mathbb{R}^6$ the Lagrangian is given by

$$L = \frac{m_1}{2}((\dot{q}^1)^2 + (\dot{q}^2)^2 + (\dot{q}^3)^2) + \frac{m_2}{2}((\dot{q}^4)^2 + (\dot{q}^5)^2 + (\dot{q}^6)^2) - V(q,\dot{q}).$$

Now $(q,\dot{q})$ can be expressed in terms of coordinates $(\hat{q},\dot{\hat{q}})$ on TQ hence L can be considered as a function on TQ.

§3. CONSERVED QUANTITIES IN THE LAGRANGE FORMALISM

DEFINITION: a function $f\colon TQ \to \mathbb{R}$ is a conserved quantity for the motion described by the Lagrangian L if

$$\frac{d}{dt} f(q(t), \dot{q}(t)) = 0$$

i.e. f is constant along the trajectories of the motion in TQ.

PROPOSITION 3.1: *the expression* $E(q,\dot{q})$ *defined by*

$$E(q,\dot{q}) = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L(q,\dot{q})$$

is independent of the choice of local coordinates (q) *on* Q, *hence* E *is a function on* TQ, *called the energy function.*

PROPOSITION 3.2: E *is a conserved quantity.*

PROOF:
$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right) =$$

(Euler-Lagrange)

$$= \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right)\dot{q}^i + \frac{\partial L}{\partial \dot{q}^i}\frac{d\dot{q}^i}{dt} - \frac{\partial L}{\partial q^i}\frac{dq^i}{dt} - \frac{\partial L}{\partial \dot{q}^i}\frac{d\dot{q}^i}{dt} = 0.$$

QED

CONSTRUCTION: if X is a vector field on Q expressed in local coordinates (q) as $X(q) = \xi^i(q)\frac{\partial}{\partial q^i}$ then we can define a lift of X to TQ as follows. Let ρ_t be the local 1-parameter group of diffeomorphisms of Q associated with X (i.e. $\left.\frac{d}{dt}\right|_{t=0} \rho_t(q) = X(q)$) then ρ_t defines a local 1-parameter group σ_t of diffeomorphisms of TQ by:

$$\sigma_t(q,\dot{q}) := (\rho_t(q), \rho_{t*}\dot{q}).$$

The lift X_+ of X to TQ is defined as the vector field on TQ associated with the flow σ_t. In local coordinates $(q,\dot{q})$ on TQ X_+ is given by

(7) $$X_+(q,\dot q) = \xi^i(q)\frac{\partial}{\partial q^i} + \frac{\partial \xi^i}{\partial q^j}(q)\dot q^j \frac{\partial}{\partial \dot q^i}.$$

Similarly ρ_t defines a flow $\hat\sigma_t \cong \rho^*_{-t}$ on T^*Q which defines a vector field X^+ on T^*Q given in local coordinates by

(8) $$X^+(q,p) = \xi^i(q)\frac{\partial}{\partial q^i} + \frac{\partial \xi^i}{\partial q^j}(q)p_i \frac{\partial}{\partial p_j}$$

where (q,p) represents the 1-form $p_i dq^i$ T^*_qQ.

<u>THEOREM</u> (E. Noether): *Let* X *be a vector field on* Q *with representation* $X(q) = \xi^i(q)\frac{\partial}{\partial q^i}$ *then:*

$$X_+L = 0 \Rightarrow f := \frac{\partial L}{\partial \dot q^i}(q,\dot q)\xi^i(q) \quad \textit{is a conserved quantity.}$$

<u>PROOF:</u> $$\frac{d}{dt}\left(\xi^i \frac{\partial L}{\partial \dot q^i}\right) = \xi^i \frac{d}{dt}\frac{\partial L}{\partial \dot q^i} + \frac{d\xi^i}{dt}\frac{\partial L}{\partial \dot q^i} =$$

$$= \xi^i \frac{\partial L}{\partial q^i} + \frac{\partial \xi^i}{\partial q^j}\frac{dq^j}{dt}\frac{\partial L}{\partial \dot q^i} = \xi^i \frac{\partial L}{\partial q^i} + \frac{\partial \xi^i}{\partial q^j}\dot q^j \frac{\partial L}{\partial \dot q^i} = X_+L = 0.$$

QED

<u>COROLLARY:</u> *if* ρ_t *is a 1-parameter group of diffeomorphisms of* Q *such that the Lagrangian* L *is invariant under the action of* ρ_t *in the sense* $X_+L = 0$ *where* $X = \xi^i \frac{\partial}{\partial q^i} = \frac{d}{dt}\Big|_{t=0}\rho_t$, *then there exists a conversed quantity* $f = \xi^i \frac{\partial L}{\partial \dot q^i}$.

§4. THE HAMILTON FORMALISM

If a Lagrangian L does not depend on a coordinate q^i then $p_i := \frac{\partial L}{\partial \dot{q}^i}$ is a conserved quantity as can be seen from Noether's theorem (or directly from the Euler-Lagrange equations). Because such a situation frequently occurs in physics, one studies the functions $p_i = \frac{\partial L}{\partial \dot{q}^i}$ which are called (generalized) momenta associated to the coordinates q^i.

<u>EXAMPLE 1</u>: $\frac{\partial L}{\partial q^1} = \frac{\partial L}{\partial q^2} = 0 \Rightarrow p_1 = m\frac{dq^1}{dt}$ and $p_2 = m\frac{dq^2}{dt}$ are conserved quantities.

<u>EXAMPLES 2</u>: $\frac{\partial L}{\partial \phi} = 0 \Rightarrow p_\phi = \frac{\partial L}{\partial \dot{\phi}} = m(q^2\dot{q}^1 - q^1\dot{q}^2)$ is conserved; p_ϕ is (a component from) the angular momentum vector $\vec{p}_{ang} = m\vec{q}\wedge\frac{d\vec{q}}{dt}$.

The definition of p_i is dependent on the choice of the (local) coordinate system and in fact if $(\hat{q})$ is a different coordinate system on Q we have:

$$\hat{p}_i = \frac{\partial L}{\partial \dot{\hat{q}}^i} = \frac{\partial L}{\partial \dot{q}^j}\frac{\partial \dot{q}^j}{\partial \dot{\hat{q}}^i} = p_j \frac{\partial q^j}{\partial \hat{q}^i}. \tag{9}$$

Comparison of this formula with the dependence of coordinates on T_q^*Q, associated to coordinates on Q, shows that the functions p_i behave as coordinates on the fibres of T^*Q. If we formalize this idea, we get the *Legendre map* associated with the Lagrangian L:

$$FL\colon TQ \to T^*Q,\ (q,\dot{q}) \longmapsto \left(q, p = \frac{\partial L}{\partial \dot{q}}\right) \tag{10}$$

where now the p_i are coordinates on T_q^*Q : (q,p) "is" the 1-form $p_i dq^i$. According to (9) this definition is independent of the coordinate system and hence well-defined on TQ.

If we want to describe formula (10) in a coordinate free way, we have to make some preliminary remarks. First: if V is a vector space, then we can identify TV and $V \times V$; the identification is given by $(v,w) \in V \times V$

"is" the tangent vector at $v \in V$ of the curve $\phi: \mathbb{R} \to V$, $t \longmapsto v + tw$. Furthermore, if f is a function on V then df is a function on TV, more precisely

$$df\big|_v : T_vV \to \mathbb{R}, \qquad w \longmapsto df\big|_v w.$$

But we identified T_vV with V hence for $v \in V$ we have $df\big|_v : V \to \mathbb{R}$ or $df\big|_v \in V^*$.

If we now turn our attention to TQ with a Lagrangian $L: TQ \to Q$ we have for $q \in Q$ and $L_q = L\big|_{T_qQ}$: $L_q : T_qQ \to \mathbb{R}$; T_qQ is a vector space hence we can apply our results:

$$v \in T_qQ \Rightarrow dL_q\big|_v \in (T_qQ)^* = T_q^*Q.$$

(N.B. the d of $dL_q\big|_v$ is the d-operator in T_qQ, NOT the d-operator in Q !) This formula now defines a map

$$FL: TQ \to T^*Q, \quad (q,v) \longmapsto (q, dL_q\big|_v)$$

which is given in local coordinates by (10).

REMARK: FL is a map which preserves fibres (hence a fibre bundle map) but in general it is not linear, so it is *not* a vector bundle map.

REMARK: FL is sometimes called the fibre derivative of L because in the coordinate independent definition it uses obviously the derivative dL_q only along the fibres.

REMARK: there exists a more geometrical interpretation of the Legendre transformation FL which is given in detail by [Arnold].

In these notes we will restrict ourselves to the case where FL is a bijection (diffeomorphism) in which case L is called a hyperregular Lagrangian

(by [Abraham & Marsden]). The case FL not bijective will be discussed by M. Bergvelt in the setting of constrained hamiltonian systems.
FL bijective implies $FL^{-1}: T^*Q \to TQ$ is a fibre preserving map and, because there exists a natural isomorphism between TQ and $T^{**}Q$ (we only consider finite dimensional manifolds), we can ask: does there exist a function H on T^*Q such that FL^{-1} is the fibre derivative of H ?

PROPOSITION 4.1: *the fibre derivative of* $H := E \circ FL^{-1}$ *is equal to* $FL^{-1}: FL^{-1} = FH$.

PROOF: we use local coordinates (q,p) on T^*Q and $(q,\dot{q})$ on TQ ; by FL^{-1} the coordinates $\dot{q}^i$ are functions of (q,p), $\dot{q}^i = \dot{q}^i(q,p)$ and H is defined by

$$H(q,p) = p_i\dot{q}^i(q,p) - L(q,\dot{q}(q,p))$$

where we used that $\frac{\partial L}{\partial \dot{q}^i}(q,\dot{q}(q,p)) = p_i$ by definition of FL. Now we have:

$$\frac{\partial H}{\partial p_i} = \dot{q}^i(q,p) + p_j\frac{\partial \dot{q}^j}{\partial p_i} - \frac{\partial L}{\partial \dot{q}^j}\frac{\partial \dot{q}^j}{\partial p_i} = \dot{q}^i(q,p).$$

QED

REMARK: $H := E \circ FL^{-1}$ is called the hamiltonian of the physical system; it is also called the function associated to L by the Legendre transformation.

REMARK: if $H^1: T^*Q \to \mathbb{R}$ is another function such that $FH^1 = FL^{-1}$ then one can (easily) show that $H - H^1$ is a function on Q or more precisely $\exists h: Q \to \mathbb{R} : H - H^1 = h \circ \pi$ where $\pi: T^*Q \to Q$ denotes the bundle projection.

The importance of the special choice of H is expressed by the following

PROPOSITION 4.2: $(q(t), \dot{q}(t))$ *is a solution of the Euler-Lagrange equations iff* $(q(t), p(t)) = FL(q(t), \dot{q}(t))$ *is a solution of*

$$(11) \qquad \frac{dq^i}{dt}(t) = \frac{\partial H}{\partial p_i}(q(t),p(t)), \quad \frac{dp_i}{dt}(t) = -\frac{\partial H}{\partial q^i}(q(t),p(t)).$$

PROOF: [$\leftarrow$] $\frac{\partial H}{\partial p_i} = \dot{q}^i$ as seen in the proof of proposition 4.1;

$$\frac{\partial H}{\partial q^i} = p_j \frac{\partial \dot{q}^j}{\partial q^i} - \frac{\partial L}{\partial q^i} - \frac{\partial L}{\partial \dot{q}^j}\frac{\partial \dot{q}^j}{\partial q^i} = -\frac{\partial L}{\partial q^i}$$

(N.B. $\dot{q} = \dot{q}(q,p)$) hence (11) $\Rightarrow$ $\frac{dq}{dt} = \dot{q}$, $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}$. The reverse implication is proved analogous. QED

EXAMPLE 1: $H(q,p) = \frac{1}{2m}(p_1^2+p_2^2+p_3^2) + mg\,q^3$

EXAMPLE 2: $H(q,p) = \frac{1}{2m}(p_1^2+p_2^2+p_3^2) - m\gamma r^{-1}$

EXAMPLE 3: $H(q,p) = \frac{1}{2m}(\vec{p} - e\vec{A}(q))^2 + e\phi(q)$

PROPOSITION 4.3: *the equations* (11) *are invariant under changes of the coordinates* (q) *on* Q.

PROOF: we apply proposition 4.2 and proposition 2.1 (invariance of the Euler Lagrange equations). QED

We finish this section by the explanation of its title: When one describes a physical system by the equations (11) (Hamilton's equations) on T^*Q with the hamiltonian function H, then one says: we use the hamilton formalism.

§5. CONSERVED QUANTITIES IN THE HAMILTON FORMALISM

THEOREM (E. Noether): *if* $X = \xi^i(q)\frac{\partial}{\partial q^i}$ *is a vector field on* Q *satisfying*

$$X^+H = 0$$

then $\xi^i(q)p_i$ *is a conserved quantity.*

PROOF: remark first that $\xi^i(q)p_i$ is a correctly defined function on T^*Q (independent of the local coordinates); the actual proof that $\xi^i(q)p_i$ is conserved is analogous to Noether's theorem in the Lagrange formalism.

QED

We see that Noether's theorem in this form generates conserved quantities of a special kind: linear in the momenta p, and we will try to generalize Noether's theorem to a larger class of functions. However, to do so, we need an intermezzo concerning Poisson brackets.

DEFINITION: if f and g are functions on T^*Q then their Poisson bracket $[f,g]$ is the function on T^*Q defined by

$$[f,g](q,p) = \frac{\partial f}{\partial p_i}(q,p)\frac{\partial g}{\partial q^i}(q,p) - \frac{\partial f}{\partial q^i}(q,p)\frac{\partial g}{\partial p_i}(q,p).$$

PROPOSITION 5.1: *the definition of* $[f,g]$ *is independent of the local coordinates on* Q, *i.e.* $[f,g]$ *is a well-defined function on* T^*Q.

PROPOSITION 5.2: *if* f, g *and* h *are functions on* T^*Q *then the Poisson brackets satisfy the relations*

(i)	$[f,g] = -[g,f]$	*(antisymmetry)*
(ii)	$[f+g,h] = [f,h] + [g,h]$	*(linearity)*
(iii)	$[f\cdot g,h] = f\cdot[g,h] + [f,h]\cdot g$	*(Leibnitz-rule)*
(iv)	$[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0$	*(Jacobi-identity)*

COROLLARY: *the set of all functions on* T^*Q *together with the operations addition and Poisson bracket form a Lie algebra.*

DEFINITION: if f is a function on T^*Q, then the associated hamilton vector field X_f on T^*Q is defined by

$$X_f\Big|_{(q,p)} = \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i}\Big|_{(q,p)} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i}\Big|_{(q,p)} .$$

Again this definition is independent of the local coordinates (q) on Q (with associated coordinates (p) on T_q^*Q).

PROPOSITION 5.3: *the solutions of Hamiltons equations* (11) *are the integral curves of* X_H.

After this intermezzo we return to the conserved quantities:

THEOREM (Noether, generalized): $f: T^*Q \to \mathbb{R}$ *is a conserved quantity iff* $[H,f] = 0$ *iff* $X_f H = 0$.

PROOF: $$\frac{d}{dt} f(q(t), p(t)) = \frac{\partial f}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}$$

$$= \frac{\partial f}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q^i} = [H,f] = -X_f H .$$

QED

REMARK: if $f(q,p) = \xi^i(q)p_i$ then $X_f = X^+$ where X is the vector field $\xi^i(q)\frac{\partial}{\partial q^i}$ on Q hence "both" theorems concur on functions linear in p.

§6. CANONICAL TRANSFORMATIONS

In this section we will derive a class of transformations (diffeomorphisms) of T^*Q for which the hamilton equations (11) are invariant. In general this class will be much larger than the class of diffeomorphisms of Q with their associated diffeomorphisms of T^*Q.

<u>DEFINITION</u>: we define a 1-form θ_o on T^*Q (i.e. θ_o is a section of $T^*(T^*Q))$ as follows: denote by $\pi: T^*Q \to Q$ the bundle projection and let α be a point of T^*Q then $q := \pi(\alpha)$ and $\alpha \in T^*_qQ$ (i.e. α is a 1-form on T_qQ). We define θ_o in α by its value on tangent vectors: if $\tau \in T_\alpha(T^*Q)$ then

(12) $$\theta_o\big|_\alpha(\tau) := \alpha(\pi_*\tau).$$

This 1-form θ_o is called the canonical 1-form on T^*Q.

<u>PROPOSITION 6.1</u>: *in local coordinates* (q) *on* Q *and associated coordinates* (p) *in* T^*Q θ_o *is given by*

$$\theta_o = p_i dq^i$$

<u>PROOF</u>: let (q^i_o, p_{io}) be the coordinates of α, i.e. $\alpha = p_{io}dq^i\big|_q$ as 1-form on T_qQ; let τ be the arbitrary tangent vector $\tau = \tau^i \frac{\partial}{\partial q^i} + \tau_j \frac{\partial}{\partial p_j}$ then $\pi_*\tau = \tau^i \frac{\partial}{\partial q^i}$ and $\theta_o\big|_\alpha(\tau) = \tau^i p_{io}$.
Now p_{io} is the value of the (coordinate) function p_i at the point α and τ^i is the value of dq^i on τ hence $\theta_o\big|_\alpha(\tau) = (p_i dq^i)\big|_\alpha(\tau)$.

QED

<u>Nota Bene</u>: the symbol dq^i in $\alpha = p_{io}dq^i\big|_q$ is quite different from the symbol dq^i in $\theta_o\big|_\alpha = p_i dq^i\big|_\alpha$: in the first case, the d-operator is the exterior derivative on Q, in the second it is the exterior derivative on

T^*Q. In fact, the functions q^i are (coordinate) functions on Q and we have used the shorthand q^i also as (coordinate) functions on T^*Q where we should have written $q^i \circ \pi$ (hence the correct expression of θ_o is given by $\theta_o = p_i\, d(q^i \circ \pi)$).

REMARK: we did not motivate the definition of θ_o; however, historically there is a very strong motivation to introduce θ_o inspired by a close analogy between classical mechanics in the hamilton formalism and geometrical optics (see for instance [Arnold]). In fact most of the definitions given in this section are motivated by this analogy.

DEFINITION: on $R = T^*Q \times \mathbb{R}$ (called the enlarged phase space) we define a 1-form α by

$$\alpha = \pi_1^* \theta_o - (H \circ \pi_1) \pi_2^* dt$$

where $\pi_1 : R \to T^*Q$ and $\pi_2 : R \to \mathbb{R}$ are the projections and t the standard coordinate on $\mathbb{R}$. In local coordinates (q,p,t) α is given by

$$\alpha = p d_q - H(q,p) dt.$$

DEFINITION: the distribution K_α on R is defined by

$$K_{\alpha,r} = \{ \xi \in T_r R \mid i_\xi d\alpha = 0 \in T_r^* R \} \subset T_r R.$$

PROPOSITION 6.2: $\forall r \in R : \dim K_{\alpha,r} = 1$.

PROOF: consider the point $r = (q,p,t)$ in local coordinates and the matrix A_{ij} defined by $A_{ij} = d\alpha(v_i, v_j)$, $i,j = 1, \ldots, 2n+1$ where (v_i) is the basis of $T_r R$ associated to the local coordinates (q,p,t) then we have

$$(13)\qquad A_{ij} = \left(\begin{array}{c:c:c} \emptyset & -I & \frac{\partial H}{\partial p} \\ \hdashline I & \emptyset & \frac{\partial H}{\partial q} \\ \hdashline -\frac{\partial H}{\partial p} & -\frac{\partial H}{\partial q} & 0 \end{array}\right) \begin{array}{c} \partial_q \\ \partial_p \\ \partial_t \end{array}$$

$$\qquad\qquad\qquad \partial_q \quad \partial_p \quad \partial_t$$

and $\text{rang}(A_{ij}) = 2n \Rightarrow \dim\ker A_{ij} = 1$.

QED

<u>COROLLARY/DEFINITION</u>: K_α is a 1-dimensional distribution hence integrable; the integral curves of K_α called the characteristics of α.

<u>PROPOSITION 6.3</u>: *if* L *is a characteristic of* α *then* $\pi_2: L \to \mathbb{R}$ *is injective; if we parametrize* L *by the coordinate* t, $L(t) = (q(t), p(t), t) \in L$ *then* $(q(t), p(t))$ *satisfy hamiltons equations*

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}.$$

<u>PROOF</u>: according to formula (13) the vector (field) X_α given by

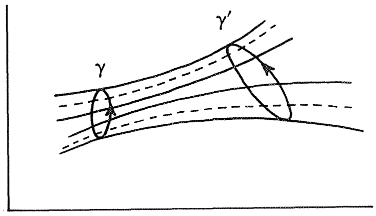
$$X_\alpha = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \partial_t$$

is an element of $K_\alpha = \ker A$. Because the coefficient of ∂_t is 1, $\pi_2: L \quad \mathbb{R}$ is injective and can be integrated to give t as a parameter along L. The other components of X_α now prove the second claim. QED

<u>DEFINITION</u>: G_t is the local 1-parameter group of diffeomorphisms of R associated to the vector field X_α (as defined in the proof of proposition 6.3). We define g_t as the local 1-parameter group of diffeomorphisms associated to X_H (see §5) or in other words: g_t defines the time-evolution of the physical system described by the hamiltonian H.

<u>COROLLARY</u>: $\pi_1{*}X_\alpha = X_H$ *and* $G_t(m,t') = (g_t(m), t'+t)$. *Furthermore if* $r \in L$ *is an arbitrary point of a characteristic* L *of* α *then* $L = \{G_t r\}$ *where* t *takes all allowed values.*

<u>DEFINITION</u>: suppose γ is a closed curve in R (i.e. $\gamma: S^1 \to R$) then we consider all characteristics of α which intersect γ. These characteristics form a so-called tube of characteristics and we say that a closed curve γ' in R encloses the same tube of characteristics if there exists a function $\tau: S^1 \to \mathbb{R}$ such that $\gamma'(p) = G_{\tau(p)}\gamma(p)$ or in words: the characteristic which intersects γ in $\gamma(p)$ intersects γ' a time $\tau(p)$ "later".



<u>PROPOSITION 6.4</u>: *if* γ *and* γ' *enclose the same tube of characteristics then*

$$\oint_\gamma \alpha = \oint_{\gamma'} \alpha$$

<u>PROOF</u>: define the surface σ as consisting of those points of the tube of characteristics lying between γ and γ' then $\partial\sigma = \gamma - \gamma'$ and we have by Stokes' Lemma:

$$\int_{\gamma-\gamma'} \alpha = \int_{\partial\sigma} \alpha = \int_\sigma d\alpha = 0$$

where the last equality follows because σ consists of characteristics of α (hence for each pair of independent tangent vectors v_1, v_2 of σ we have $d\alpha(v_1,v_2) = 0$). QED

COROLLARY: *if* γ *is a closed loop in* T^*Q *then*

$$\oint_\gamma \theta_o = \oint_{g_t\gamma} \theta_o \quad \textit{or} \quad \oint_\gamma pdq = \oint_{g_t\gamma} pdq .$$

COROLLARY: $g_t^* d\theta_o = d\theta_o$ *or in local coordinates: the 2-form* $dp_i \wedge dq^i$ *is invariant under the flow of* X_H.

PROOF: let σ be any surface in T^*Q such that $\partial\sigma = \gamma$ a single closed loop then

$$\int_\sigma d\theta_o = \int_\gamma \theta_o = \int_{g_t\gamma} \theta_o = \int_\gamma g_t^*\theta_o = \int_\sigma g_t^* d\theta_o .$$

QED

COROLLARY: *the volume element* $\varepsilon = dp_1 \wedge \ldots \wedge dp_n \wedge dq^1 \wedge \ldots \wedge dq^n$ *on* T^*Q *is conserved under* g_t.

PROOF: $\varepsilon = (n!)^{-1} \cdot (-1)^{\frac{1}{2}n(n-1)} (d\theta_o)^n$.

QED

At this point we have gathered enough results to find/define a class of diffeomorphisms under which the hamilton equations are invariant.

DEFINITION: a diffeomorphism $g\colon T^*Q \to T^*Q$ is called a canonical transformation if $g^* d\theta_o = d\theta_o$.

PROPOSITION 6.5: *suppose* g *is a canonical transformation of* T^*Q, *and suppose* q^i *are local coordinates on* Q *with associated coordinates* p_i *on* T^*Q *then the system* (Q^i, P_i) *defined by* $Q^i := q^i \circ g$, $P_i := p_i \circ g$ *also is a system of coordinates on* T^*Q *and we have: the equations of motion on* T^*Q *are given by:*

$$\frac{dQ^i}{dt} = \frac{\partial H}{\partial P_i} , \quad \frac{dP_i}{dt} = -\frac{\partial H}{\partial Q^i} .$$

PROOF: $g^* d\theta_o = d\theta_o \Rightarrow dp_i \wedge dq^i = dP_i \wedge dQ^i$ hence if $\beta := p_i\, dq^i - P_i\, dQ^i$ then $d\beta = 0$ and we have $p_i\, dq^i - H\, dt = P_i\, dQ^i - H\, dt + \beta$.

Now because $d\beta = 0$, the characteristics of $p_i\, dq^i - H\, dt$ are equal to the characteristics of $P_i\, dQ^i - H\, dt$ and we then apply proposition 6.3. QED

COROLLARY: *the form of the hamilton equations* (11) *is invariant under canonical transformations of* T^*Q.

PROPOSITION 6.6: *if* h *is a diffeomorphism of* Q *then* $g := h^*$ *is a canonical transformation of* T^*Q.

PROOF:
$$(g^*\theta_o)\big|_\alpha(\tau) = \theta_o\big|_{g\alpha}(g_*\tau) = (g\alpha)(\pi_* g_* \tau)$$
$$= (h^*\alpha)((\pi \circ g)_* \tau) = (h^*\alpha)((h^{-1} \circ \pi)_* \tau) = (h^*\alpha)(h_*^{-1} \pi_* \tau)$$
$$= \alpha(\pi_* \tau) = \theta_o\big|_\alpha(\tau) \Rightarrow g^*\theta_o = \theta_o \Rightarrow g^* d\theta_o = d\theta_o .$$
QED

REMARK: from the proof of proposition 6.6 we see that $g \equiv h^*$ satisfies the condition $g^*\theta_o = \theta_o$ which is a stronger condition that $g^* d\theta_o = d\theta_o$!

§7. SYMPLECTIC GEOMETRY

In the previous two sections all definitions were given in (local) coordinates although we showed that they were independent of these coordinates, e.g. hamiltons equations, Poisson brackets and hamiltonian vector fields. The question arises which properties are due to special choices of our coordinate systems, and which properties are valid in general, i.e. we ask ourselves: what is the essential structure needed to prove "all" our propositions?

The answer is given by symplectic geometry; it turns out that very little structure is needed: only a manifold and a two-form satisfying certain conditions (a so-called symplectic manifold). In this section we will outline this "construction" and we will reformulate the previous two sections in a coordinate-free way, showing that the structure of a symplectic manifold indeed is the only essential ingredient.

Nota Bene: the context in which we will reformulate the previous sections is slightly more general than the context of these sections. The reader should be aware of this fact when he thinks some proofs are superfluous.

DEFINITION: a symplectic manifold is a pair (M,ω) in which M is a manifold and ω a closed nondegenerate 2-form on M, i.e.

$$d\omega = 0$$

$$\phi: T_mM \to T_m^*M, \quad X \longmapsto i_X\omega \quad \text{is an isomorphism.}$$

PROPOSITION 7.1: *the dimension of a symplectic manifold* (M,ω) *is even:* $\dim M = 2n$.

PROOF: consider a basis (e_i) of T_mM and the matrix $A_{ij} := \omega(e_i,e_j)$ then A is skew symmetric and:

$$\det A = \det(-A^T) = (-1)^{\dim M}\det(A^T) = (-1)^{\dim M}\det A$$

hence if $\dim A$ is odd then $\det A = 0$ and ϕ cannot be an isomorphism.

EXAMPLES: 1. if Q is a manifold then (T^*Q, ω_o) is a symplectic manifold when $\omega_o = d\theta_o$; ω_o is called the canonical 2-form on T^*Q.

2. $M = \mathbb{P}^n\mathbb{C}$ the complex projective space of complex dimension n carries a natural symplectic structure. If $(z_o:\ldots:z_n)$ are homogeneous coordinates on M then we have local complex coordinates $(w_1,\ldots,w_n) = \left(\frac{z_o}{z_i},\ldots,\frac{z_{i-1}}{z_i},\frac{z_{i+1}}{z_i},\ldots,\frac{z_n}{z_i}\right)$ on $U_i = \{z_i \neq 0\}$. The real and imaginary parts of w_i form a set of $2n$ real local coordinates and in terms of these coordinates the symplectic form ω is given by

$$\omega = i\frac{\partial^2 f}{\partial w_a \partial \bar{w}_b}\, dw_a \wedge d\bar{w}_b$$

where $f(w_1,\ldots,w_n) = \ln(1 + \sum_{a=1}^{n} w_a\bar{w}_a)$.

3. As a special case of 2) we have $S^2 \cong \mathbb{P}^1\mathbb{C}$ and the symplectic form ω is given by:

$$\omega = \frac{i}{(1+w\bar{w})^2}\, dw \wedge d\bar{w} \quad \text{on } U_o \text{ with coordinate } w.$$

In real coordinates (p,q), $w = p + iq$ we have

$$\omega = \frac{2}{(p^2+q^2+1)^2}\, dp \wedge dq.$$

If we use the identification of coordinates on S^2:

$$(p,q) \longleftrightarrow \begin{pmatrix} 2p/(1+p^2+q^2) \\ 2q/(1+p^2+q^2) \\ (p^2+q^2-1)/(1+p^2+q^2) \end{pmatrix} = \begin{pmatrix} \sin\theta\,\sin\phi \\ \sin\theta\,\cos\phi \\ \cos\theta \end{pmatrix} \longleftrightarrow (\theta,\phi)$$

then ω is given by

$$\omega = \tfrac{1}{2}\sin\theta\; d\theta \wedge d\phi$$

which is the usual volume element on S^2 times $\frac{1}{2}$.

In proposition 6.5 we saw that the canonical symplectic form ω_o on T^*Q is an important object in classical mechanics. The famous theorem of Darboux states that locally every symplectic form ω looks like ω_o.

<u>THEOREM</u> (Darboux): *let* (M,ω) *be a* 2n-*dimensional symplectic manifold and* $m \in M$, *then there exist local coordinates* $(q^i,p_i)_{i=1}^n$ *around* m *such that* $\omega = dp_i \wedge dq^i$ *locally.*

The proof we will give of this theorem is almost verbatim taken from [Woodhouse] and is due to Moser and Weinstein; it depends upon the following lemma of Moser.

<u>LEMMA</u>: let ω and σ be two closed nondegenerate 2-forms on M; if for some $m \in M$ $\omega_m = \sigma_m$ then there exist neighbourhoods U and V of m and a diffeomorphism $\rho: U \to V$ such that $\rho(m) = m$ and $\rho^*\sigma = \omega$.

<u>PROOF</u>: $d(\sigma-\omega) = 0 \Rightarrow \exists$ a neighbourhood W of m and a 1-form α on W: $d\alpha = \sigma - \omega \wedge \alpha_m = 0$. Define $N = W \times [0,1]$ and denote by $pr: N \to W$ the projection on the first factor; we then define a 2-form Ω on N by:

$$\Omega = pr^*\omega + t\cdot pr^*(\sigma-\omega) + dt \wedge pr^*\alpha.$$

For each $t \in [0,1]$ we define $i_t: W \to N$ by $i_t(w) = (w,t)$ and we define $\Omega_t = i_t^*\Omega = \omega + t(\sigma-\omega)$. If we choose W sufficiently small, then Ω_t is non-degenerate for all $t \in [0,1]$. Using the same technique as in the proof of propositions 6.2, 6.3 one can show that there exists a unique vector field X on N such that

$$dt(X) = 1, \quad i_X\Omega = 0.$$

Let ϕ_t be the flow associated to X and for each $x \in W$ let $t \longmapsto \phi_x(t) = \phi_t(x,o)$ be the integral curve of X through $(x,o) \in N$. Then $dt(X) = 1$ implies $\phi_x(t) \in W \times \{t\}$ for all values of t for which ϕ_x is defined. Since $\Omega\big|_{\{m\} \times [0,1]} = pr^*\omega_m$ $(\omega_m = \sigma_m \wedge \alpha_m = 0)$ if follows that $\phi_m(t) = (m,t)$ for all $t \in [0,1]$ hence there exists a neighbourhood U of m such that $\phi_x(t)$ is defined for all $(x,t) \in U \times [0,1]$.

Now define $\rho : U \to V = \rho(U)$ by $\phi_x(1) = (\rho(x), 1)$ and notice that $L_X\Omega = i_X d\Omega + di_X\Omega = 0$ and $\phi_1 \circ i_o = i_1 \circ \rho \Rightarrow i_o^* \circ \phi_1^*\Omega = \rho^* \circ i_1^*\Omega \Rightarrow$ $\Rightarrow i_1^*\Omega = \rho^* i_1^*\Omega$ (because $L_X\Omega = 0$) $\Rightarrow \Omega_o = \rho^*\Omega_1$. From the definition of Ω_t it follows that $\rho^*\sigma = \omega$. QED

<u>PROOF of Darboux's theorem</u>: consider the matrix $\omega_{ij} = \omega(e_i, e_j)$ where (e_i) is a basis of T_mM associated to some coordinate system around m. It is elementary linear algebra to show that there exist local coordinates $(r^i, s_i)_{i=1}^n$ around m such that

$$\omega_{ij} = \begin{pmatrix} \emptyset & -I \\ I & \emptyset \end{pmatrix} \quad \text{or (equivalently)}$$

$$\omega_m = ds_i \wedge dr^i .$$

Now define $\sigma = ds_i \wedge dr^i$ around m then $\sigma_m = \omega_m$ and we can apply Moser's lemma. Let $p_i = s_i \circ \rho$, $q^i = r^i \circ \rho$ then (q^i, p_i) are local coordinates around m and $\omega = \rho^*\sigma = dp_i \wedge dq^i$. QED

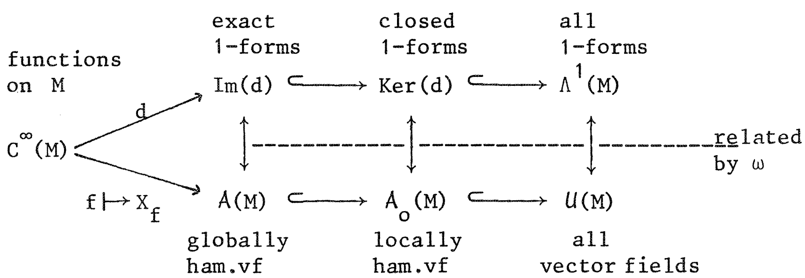
<u>DEFINITION</u>: a coordinate system (q^i, p_i) on a symplectic manifold (M,ω) is called a canonical coordinate system if $\omega = dp_i \wedge dq^i$ on the domain of (q^i, p_i).

<u>COROLLARY</u>: *every point on a symplectic manifold has a neighbourhood on which there exist canonical coordinates.*

We now turn our attention to hamilton vector fields: on a symplectic manifold (M,ω), ω_m defines an isomorphism between T_mM and T_m^*M hence ω defines an isomorphism between vector fields and 1-forms by:

$$\xi \leftrightarrow \alpha \iff i_\xi \omega + \alpha = 0.$$

<u>DEFINITION</u>: a vector field ξ on M is called a locally hamiltonian vector field if the associated 1-form α defined by $i_\xi \omega + \alpha = 0$ is a closed 1-form; ξ is called a (globally) hamiltonian vector field if α is exact. If ξ is a hamiltonian vector field and $i_\xi \omega + df = 0$ then we denote $\xi = X_f$ and ξ is called the hamiltonian vector field associated to the function f. We can express the relations between functions, 1-forms and locally/globally hamiltonian vector fields in the following diagram:



<u>PROPOSITION 7.2</u>: *in local canonical coordinates* (q,p) *on* (M,ω) *the vector field* X_f *is given by:*

$$X_f = \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i}.$$

<u>PROOF</u>: $\omega = dp_i \wedge dq^i$, $X_f = \alpha^i \frac{\partial}{\partial q^i} + \beta_i \frac{\partial}{\partial p_i}$, $i_{X_f}\omega + df = 0 \Rightarrow$ for any vector field Y on M, $Y = Y^i \frac{\partial}{\partial q^i} + Y_j \frac{\partial}{\partial p_j}$ we have:

$$i_{X_f}\omega(Y) = \omega(X_f, Y) = \beta_i Y^i - \alpha^i Y_i = -df(Y)$$
$$= -\left(\frac{\partial f}{\partial q^i} Y^i + \frac{\partial f}{\partial p_i} Y_i\right)$$

hence

$$\beta_i = -\frac{\partial f}{\partial q^i} \quad \text{and} \quad \alpha_i = \frac{\partial f}{\partial p_i} .$$

QED

<u>PROPOSITION 7.3</u>: *let* H *be a hamiltonian on* T^*Q *and* ω *the canonical symplectic form. Then the time evolution of the physical system described by this hamiltonian is given by the flow of the vector field* X_H *(see proposition 5.3).*

<u>PROPOSITION 7.4</u>: *if* $\xi \in U(M)$ *then* $\xi \in A_o(M) \iff L_\xi \omega = 0$.

<u>PROOF</u>: let α be defined by $i_\xi \omega + \alpha = 0$ then (using the identity $L_\xi = i_\xi d + d\, i_\xi$) we have:

$$L_\xi \omega = 0 \iff i_\xi d\omega + d\, i_\xi \omega = 0 \iff d\, i_\xi \omega = 0$$
$$\iff d\alpha = 0 \iff \xi \in A_o(M).$$

QED

<u>DEFINITION</u>: a canonical transformation ρ of a symplectic manifold (M,ω) is a diffeomorphism ρ of M such that $\rho^*\omega = \omega$.

<u>COROLLARY</u>: *the flow of a locally hamiltonian vector field consists of canonical transformations, in particular if* ρ_t *is the flow associated to* X_H *then* $\rho_t^*\omega = \omega$.

<u>PROPOSITION 7.5</u>: $\xi, \eta \in A_o(M) \Rightarrow [\xi, \eta] = X_{\omega(\xi,\eta)} \in A(M)$.

PROOF: using the identity $i_{[\xi,\eta]} = L_\xi\, i_\eta - i_\eta\, L_\xi$ we have

$$i_{[\xi,\eta]}\omega = L_\xi\, i_\eta\omega - i_\eta\, L_\xi\omega$$

$$= d\, i_\xi\, i_\eta\omega + i_\xi d\, i_\eta\omega \qquad \text{(proposition 7.4)}$$

$$= -d\omega(\xi,\eta) \qquad (\eta \in A_o(M))$$

hence $[\xi,\eta] = X_{\omega(\xi,\eta)}$.

QED

COROLLARY: $A(M)$ *is an ideal in* $A_o(M)$ *(with addition and commutator brackets)*

DEFINITION: the Poisson bracket of two functions f and g on M is defined by

$$[f,g] = \omega(X_f,X_g)$$

COROLLARY: $[X_f,X_g] = X_{[f,g]}$.

PROPOSITION 7.6: *in local canonical coordinates* (q,p) $[f,g]$ is given by:

$$[f,g] = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}.$$

PROOF: by definition of X_f we have:

$$[f,g] = \omega(X_f,X_g) = (i_{X_f}\omega)(X_g) = (-df)(X_g) = -X_g f$$

and we apply proposition 7.2. QED

PROPOSITION 7.7: *if* f,g *and* h *are functions on* M: *then the following equalities hold:*

(i)	$[f,g] = -[g,f]$	*(antisymmetry)*
(ii)	$[f+g,h] = [f,h]+[g,h]$	*(linearity)*
(iii)	$[f\cdot g,h] = f\cdot[g,h]+[f,h]\cdot g$	*(Leibnitz rule)*
(iv)	$[f,[g,h]]+[g,[h,f]]+[h,[f,g]] = 0$	*(Jacobi-identity)*

PROOF:

(i) $[f,g] = \omega(X_f,X_g) = -\omega(X_g,X_f) = -[g,f]$.

(ii) $[f+g,h] = -X_h(f+g) = -X_hf - X_hg = [f,h]+[g,h]$

(iii) $[f\cdot g,h] = -X_h(f\cdot g) = -f\,X_hg - g\,X_hf = f\cdot[g,h]+[f,h]\cdot g$.

(iv) this is a consequence of the closedness of ω!

$$\begin{aligned} 0 &= d\omega(X_f,X_g,X_h) \\ &= X_f\omega(X_g,X_h) - X_g\omega(X_f,X_h) + X_h\omega(X_f,X_g) \\ &\quad - \omega([X_f,X_g],X_h) + \omega([X_f,X_h],X_g) - \omega([X_g,X_h],X_f) \\ &= X_f[g,h] + X_g[h,f] + X_h[f,g] \\ &\quad - \omega(X_{[f,g]},X_h) - \omega(X_{[h,f]},X_g) - \omega(X_{[g,h]},X_f) \\ &= 2\{[f,[g,h]]+[g,[h,f]]+[h,[f,g]]\}. \end{aligned}$$

QED

COROLLARY: *the map* $f \longmapsto X_f$ *is a Lie algebra homomorphism from* $(C^\infty(M),+,[\ ,\ ])$ *to* $(A(M),+,[\ ,\ ])$. *The kernel of this map consists of all locally constant functions on* M, *in particular if* M *is connected then* $\ker \simeq \mathbb{R}$.

§8. REFERENCES

ABRAHAM, R.- MARSDEN, J.E., *Foundations of mechanics*. 2° rev.& enl.ed., Benjamin, 1977.

ARNOLD, V., *Mathematical methods of classical mechanics*. Springer, 1978 (GTM 60).

WOODHOUSE, N., *Geometric quantization*. Clarendon press, 1980.

The Hamiltonian Structure of Yang-Mills Theories

M.J. Bergvelt
Institute of Theoretical Physics, University of Amsterdam

INTRODUCTION

Nowadays one hardly needs an excuse to study Yang-Mills theories. In this report we will investigate pure Yang-Mills theory (i.e. gauge fields without matter) as a classical dynamical system.

It is well-known that Yang-Mills theory is, just as Maxwell theory, an example of what is called a singular dynamical system. This means that the coordinates and momenta one starts with are not independent, there are constraints imposed on them. These constraints may lead to gauge freedom and one has to be careful in determining the true degrees of freedom and in deriving the equations of motion.

Dirac [9] and Bergmann [28] around 1950 developed a formalism to treat singular systems. This formalism, while algebraically sound did not give much insight into what one is actually doing.

Since then classical mechanics has been formulated in the language of differential geometry ([2], [3], [7]) and also the Dirac-Bergmann theory can be fitted in this framework, as was shown by Gotay, Nester and Hinds [9], giving a very intuitive picture of the theory of constraints.

Many interesting physical theories are singular in the above sense. Two recent reviews of this subject are Sundermeyer [15] and Hanson, Regge and Teitelboim [42]. However, they use the Dirac-Bergmann formalism. The geometric theory can be found in [9] and in papers by Gotay and Nester [11] and Gotay [12]. See also Lichnerowitz [10] and Sniatycki [12].

We will apply in this report the Gotay, Nester and Hinds theory to the Yang-Mills system.

The results we obtain in this way (the "Reduced Phase Space" in particular) are not new, they also follow from the traditional Dirac-Bergmann method (see for example Mitter [24]). We hope, however, that the geometrical approach clarifies the usual treatment of Yang-Mills theory in the literature.

Another point is that there are indications (infinite number of conservation laws, Bäcklund transformations) that Yang-Mills theory might be completely integrable, just as the Korteweg-de Vries equation and the Sine-Gordon equation are. This phenomenon of complete integrability is understood (at least for finite dimensional systems) on a fundamental level using the geometrical formulation of mechanics, see for example Symes [43]. Also in this respect a geometric formulation of the dynamics of Yang-Mills theories seems useful.

This report is organized as follows: We try to be as self-contained as possible, so we start with explaining in Chapter 1 the modern formulation of mechanics, (as far as we need it of course, it is a vast subject [2]). We assume that the reader is familiar with the basic concepts of differential geometry, such as manifolds, tangent vectors, differential forms, Lie derivatives etc. A good elementary introduction for physicists is Schutz [31]. Mathematically more sophisticated is Abraham & Marsden [2], the notation of which we will try to follow. In Chapter 2 we summarize the Gotay-Nester-Hinds theory and we apply this in Chapter 3 to the Yang-Mills system. We find that the "true phase space" of Yang-Mills theory is the cotangent bundle of the space of inequivalent gauge potentials (orbit space). We summarize briefly in the last part of Chapter 3 what is known about orbit space and the relation with the Faddeev-Popov quantization of gauge theories.

ACKNOWLEDGEMENTS

It is a pleasure to thank Dr. E.A. de Kerf for suggesting to write this report in the first place, and for many stimulating discussions. Also I would like to thank Mrs. H.E. v.d. Beld-Smit for her patience in deciphering my handwriting to create a nice typescript This work is finally supported by the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)".

Chapter 1

CLASSICAL MECHANICS IN THE LANGUAGE OF DIFFERENTIAL GEOMETRY

1.1. Introduction

Traditionally physicists formulated classical mechanics, or phyisics in general, in terms of coordinates with respect to some reference frame. This approach tends to lead to clumsy formulae ([1], see also fig. 3.1 of [17]) and it obscures the geometric content of the theory one is studying.

For these reasons there is a trend in modern physics to formulate theories in coordinate independent way. (See for instance [25], [26], [17].) In this chapter we will show how one can do this for classical mechanics, using the language of differential geometry, [2], [3], [4], [7].

We start with the traditional formulation:
A system of n degrees of freedom is described by n generalized coordinates $(q_1,q_2,\dots q_n)$ and n generalized velocities $(\dot{q}_1,\dot{q}_2,\dots,\dot{q}_n)$. The dynamics is determined by a Lagrangian function $L(q_i,\dot{q}_i)$, which appears in the Euler-Lagrange equations, the equations of motion of the system:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \ . \tag{1.1.1}$$

This is the Lagrangian formalism. An equivalent description is given by the Hamiltonian formalism. One introduces a momentum p^i conjugate to q_i by

$$p^i = \frac{\partial L}{\partial \dot{q}_i} \tag{1.1.2}$$

and a Hamiltonian function $H(q_i,p^i)$ by

$$H(q_i,p^i) = p^i\dot{q}_i(q_j,p^j) - L(q_i,\dot{q}_i(q_j,p^j)).$$

This definition makes sense only if it is possible to express, using (1.1.2), the velocities as functions of the coordinates and the momenta. If this is not possible one speaks of a *constrained* Hamiltonian system. The study of these systems will be our major topic. In this introductory chapter, however, we will assume that there are no constraints. The equations of motion equivalent to (1.1.1) are

$$\frac{dq_i}{dt} = \frac{\partial H_i}{\partial p^i}$$
$$\frac{dp^i}{dt} = \frac{-\partial H}{\partial q_i} \tag{1.1.4}$$

These are the well-known Hamiltonian equations.

In the rest of this chapter we will rephrase this familiar theory in geometric terms.

We take *configuration space* M, the space of which the q_i's are coordinates, to be an n-dimensional differentiable manifold. (So we consider first systems with a finite number of degrees of freedom. In section 1.6 we generalize to infinite systems). *Velocity phase space*, the space of q_i's and $\dot{q}_i$'s, is then the tangent bundle TM of M; the space coordinatized by q_i's and p^i's, *phase space*, is the cotangent bundle T^*M of M. We leave it to the reader to proof that the $\dot{q}_i$'s and p^i's have the correct transformation character to make this identification with the tangent bundle and the cotangent bundle possible.

Remark that in general one needs more than one set of coordinates to cover the whole of configuration space. For instance, if one describes a particle moving on the surface of a sphere in three dimensions (so $M = S^2$) one needs at least two coordinate systems.

The Lagrangian and the Hamiltonian are functions on the tangent and cotangent bundle:

$$L: TM \to \mathbb{R}$$

$$H: T^*M \to \mathbb{R}$$

In the next section we will give the intrinsic connection between these functions and between the tangent and cotangent bundles of configuration space.

1.2. The fiber derivative.

The transformation (1.1.2) from velocities to momenta can be interpreted as a mapping:

$$FL: TM \to T^*M$$

by demanding that in local coordinates of TM and T^*M one has:

$$FL(q_i,\dot{q}_i) = (q_i, p^i = \frac{\partial L}{\partial \dot{q}_i}) \ . \qquad (1.2.1)$$

FL is called the fiber derivative of the Lagrangian.

We see that FL maps the fibres of TM into the fibres of T^*M (see fig. 1.)

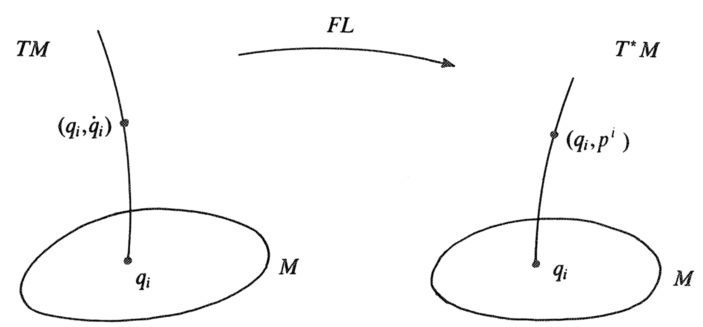


Fig. 1. The fiber derivative.

The assumption that there are no constraints, which we make in this chapter, is equivalent to the assumption that FL is bijective. In the case of constraints FL will fail to be bijective.

We can define FL implicitly in a coordinate free way as follows:

$$\langle z \mid FL(w) \rangle = \frac{d}{dt} L(w+tz)\Big|_{t=0} . \tag{1.2.2}$$

Here z and w are points of TM with the same projection on M, i.e. z and w are points of the same fiber, and one can define the sum of them, as is used in the right-hand side of (1.2.2). The bracket on the left-hand side of (1.2.2) is the contraction between an element of T^*M, FL(w) and an element of TM, z. The name fiber derivative comes from the right-hand side of (1.2.2): one takes there the partial derivative of L in some direction in the fiber.

Using the fiber derivative we define the Hamiltonian by:

$$H \circ FL(w) = \langle w \mid FL(w) \rangle - L(w) \qquad w \in TM. \tag{1.2.3}$$

We now show that in local coordinates FL defined by (1.2.2) has the form (1.2.1) and that the Hamiltonian (1.2.3) has the local form (1.1.3).

Let in local coordinates

$$z = (q_i, \dot{z}_i), \qquad w = (q_i, \dot{w}_i)$$

and

$$FL(w) = (q_i, p^i).$$

Then the right-hand side of (1.2.2) becomes

$$\frac{d}{dt} L(q_i, \dot{w}_i + t\dot{z}_i)\Big|_{t=1} = \frac{\partial L}{\partial \dot{q}_i} \dot{z}_i .$$

The left-hand side of (1.2.2) is

$$p^i \dot{z}_i .$$

Equating these results we obtain

$$p^i = \frac{\partial L}{\partial \dot{q}_i}$$

and FL is of the form (1.2.1).

In local coordinates the Hamiltonian becomes

$$H \circ FL(q_i, \dot{q}_i) = p^i \dot{q}_i - L(q_i, \dot{q}_i)$$

and this is just the form (1.1.3).

We stress that the mapping FL and the Hamiltonian are defined by (1.2.2) and (1.2.3) without reference to coordinates. We have here an intrinsic way to define FL and H.

1.3. The geometric meaning of Hamilton's equations.

In this section we will interpret Hamilton's equations (1.1.4) as the equations for the integral curves of a vector field on phase space P.

A vector field on P is a mapping $X: P \to TP$. In local coordinates a point u in P is given by

$$u = (q_i, p^i)$$

and the vector field X by

$$X_u = a^i(u) \frac{\partial}{\partial q_i} + b_i(u) \frac{\partial}{\partial p^i} \quad . \tag{1.3.1}$$

An integral curve $\ell(t)$ of the vector field X_u is a curve in P which satisfies:

$$\frac{d}{dt} \ell(t) \Big|_u = X_u(\ell(t)) \ , \tag{1.3.2}$$

(in other words: the tangent vector to $\ell(t)$ is in every point equal to the value of the vector field X in that point). In local coordinates we have

$$\ell(t) = (q_i(t), \ p^i(t))$$

and (1.3.2) reads (using (1.3.1)):

$$\frac{d}{dt}q_i(t) = a^i(q_i(t),\ p^i(t))$$
$$\frac{d}{dt}p_i(t) = b_i(q_i(t),\ p^i(t))\ . \qquad (1.3.3)$$

Comparing (1.3.3) with the Hamilton's equations (1.1.4), we see that solutions of (1.1.4), the physical trajectories on P, are integral curves of the so-called Hamiltonian vector field X_H :

$$X_H = \frac{\partial H}{\partial p^i}\frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p^i}\ . \qquad (1.3.5)$$

Denoting an integral curve of X_H through u by $\ell_H(u,t)$ we see that the time evolution of the system is given by the flow F_t of X_H :

$$F_t\colon u \to F_t(u) = \ell_H(u,t)$$
$$F_o\colon u \to F_o(u) = \ell_H(u,o) = u.$$

The Hamiltonian vector field X_H is closely related to the differential dH of the Hamiltonian. In local coordinates

$$dh = \frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial p^i}dp^i\ .$$

The components of this 1-form are, apart from ordering and some signs, precisely the components of the Hamiltonian vector field X_H .

We have defined X_H in (1.3.5) using coordinates, but we want to give an intrinsic definition. Since dH, the differential of a function, is an intrinsic object, it is useful to investigate the relationship between dH and X_H . This is the subject of the next section.

1.4. The symplectic form, coordinate dependent definition.

The relationship between dH and X_H can be put in an elegant form by introducing a 2-form ω on P, which has in local coordinates the form:

$$\omega = dp^i \wedge dq_i \; . \tag{1.4.1}$$

One can easily checks that for arbitrary vector fields Y on P :

$$\omega(X_H, Y) = -dH(Y) \; . \tag{1.4.2}$$

This gives the connection between dH and X_H that we were looking for. We can look upon (1.4.2) as an equation for the Hamiltonian vector field X_H, if the Hamiltonian H is given. But then the mapping

$$\Omega \colon TP \to T^*P$$

defined by

$$\langle Y \mid \Omega(X) \rangle = \omega(X,Y) \tag{1.4.3}$$

(with X and Y in TP) has to be surjective to insure that at least one solution of (1.4.2) exists and has to be injective to insure that not more than one solution exists. In short: Ω has to be bijective. If this is the case ω is called symplectic. (In general any closed 2-form is called symplectic if the associated mapping (1.4.3) is bijective).

We show that ω is symplectic by choosing a local basis for TP :

$$\left(\frac{\partial}{\partial q_i} \, , \, \frac{\partial}{\partial p^i}\right)$$

and using (1.4.3) and (1.4.1) to obtain

$$\begin{aligned} \Omega\left(\frac{\partial}{\partial q_i}\right) &= -dp^i \\ \Omega\left(\frac{\partial}{\partial p^i}\right) &= dq^i \end{aligned} \tag{1.4.4}$$

The inverse mapping is

$$\Omega^{-1}(dq_i) = \frac{\partial}{\partial p^i}$$
$$\Omega^{-1}(dp^i) = -\frac{\partial}{\partial q_i} \qquad (1.4.5)$$

and Ω is a bijection and ω symplectic.

Another definition, equivalent to (1.4.3), of Ω is given by

$$\Omega(X) = i_X\omega$$

where i_X is the contraction, from the left, of X with a form.

With this notation we write for (1.4.2)

$$i_{X_H}\omega = -dH . \qquad (1.4.6)$$

We refer to equation (1.4.6), following the practice in the literature [9], as the equations of motion of our system (even if it would perhaps be more logical to use this name for the equation (1.3.2) for the integral curves of X_H).

Since we have used coordinates in (1.4.1) to define the symplectic form ω the equations of motion (1.4.6) are still coordinate dependent. In the next section we give an intrinsic definition of ω.

1.5. The symplectic form, intrinsic definition.

We first define on phase space a canonical 1-form θ, the Liouville form, and then define the symplectic form ω as the exterior derivative of θ.

Consider the cotangent bundle $P = T^*M$ of M. The bundle projection π is a mapping from P to the base space M:

$$\pi: P \to M .$$

The derivative π_* of this map is a mapping from the tangent bundle TP on the tangent bundle TM

$$\pi_*: TP \to TM .$$

Let u be a point in P, i.e. u is a 1-form on M, and Y a vector in TP. We then define the canonical 1-form θ in T^*P by:

$$\langle \theta \mid Y_u \rangle = \langle u \mid \pi_* Y_u \rangle . \tag{1.5.1}$$

In local coordinates we have

$$\begin{aligned} u &= (q_i, p^i) \\ \pi&: u \longmapsto (q_i) \\ Y &= a_i \frac{\partial}{\partial q_i} + b^i \frac{\partial}{\partial p^i} \\ \pi_* Y_u &= a_i \frac{\partial}{\partial q^i} \quad . \end{aligned} \tag{1.5.2}$$

The 1-form u in T^*M acting on TM reads

$$u = p^i dq_i .$$

We obtain for the right-hand side of (1.5.1):

$$a_i p^i .$$

This leads to

$$\theta = p^i dq_i . \tag{1.5.3}$$

Note, however, that θ is defined by (1.5.1) without reference to coordinates.

From the canonical 1-form θ we define the 2-form ω by

$$\omega = d\theta . \tag{1.5.4}$$

In local coordinates

$$\omega_u = dp^i \wedge dq_i$$

in accordance with (1.4.1).

So we have in (1.5.4) an intrinsic definition of the symplectic form ω and consequently an intrinsic formulation of mechanics. We briefly summarize what we have done to achieve this.

We started with a finite dimensional differentiable manifold M as configuration space. On velocity phase space TM there was a Lagrangian function L. Using L we defined in (1.2.2) the fiber derivative FL, which was in its turn used to define H (1.2.3). We assumed that FL was bijective. On phase space we defined the canonical 1-form θ (1.5.1) and the symplectic form ω. The Hamiltonian vector field was obtained from (1.4.6). The dynamical trajectories of the system are integral curves (1.3.2) of X_H.

We stress that we never needed to introduce coordinates to describe the general framework of mechanics.

1.6. Infinite dimensional systems. [5], [6]

The formulation of Hamiltonian mechanics given in the preceedings sections applies to systems with a finite number of degrees of freedom. In this case the configuration spaces, phase spaces etc. are finite dimensional. We want to study field theories, where the configuration spaces (the spaces of initial conditions) are function spaces. To include these theories we have to generalize the formalism.

The basic mathematical notion we need is that of *a Banach-manifold*, which is a straightforward extension of the idea of a manifold to infinite dimensions. Locally a Banach manifold is homeomorphic to a (possibly infinite dimensional) Banach space E. Thus we have an atlas of charts of M, consisting of pairs (U_i,ϕ_i) with U_i open in M and

$$\phi_i : U_i \to E$$

a homeomorphism to some fixed Banach space E, such that if

$$U_i \cap U_j \neq \emptyset$$

the transition function $\phi_i \circ \phi_j^{-1}$ is continuous, of order C^k, analytic etc. depending on whether one studies a C^o, C^k or C^ω Banach manifold. In the standard way one defines related structures such as the tangent bundle, 1-forms etc.

Remark that finite dimensional manifolds are also Banach manifolds.

We can now apply the geometric theory of dynamics to the infinite dimensional case. To explain how one deals with Banach manifolds we define and calculate the canonical 1-form θ and the symplectic 2-form ω for a system the configuration space M of which is a Banach manifold.

We first give a list of the manifolds needed for the formulation of Hamiltonian mechanics

M	q	U	x
TM	V_q	$U \times E$	(x,e)
$P = T^*M$	p_q	$U \times E^*$	(x,α)
TP	Yp_q	$U \times E^* \times E \times E^*$	(x,α, e,β)
T^*P	θ_{p_q}	$U \times E^* \times E^* \times E^{**}$	(x,α,γ, f)

Table. 1.

The first column contains the various manifolds, the second gives typical points in these manifolds. The third one gives a natural coordinate neighbourhood of that point and the last column gives the coordinates in a natural chart. We use Latin characters for elements of E and Greek ones for elements of E^*. We give some comments to explain the table. M is a Banach manifold modelled on E, so every point q of M has a coordinate neighbourhood U which is homeomorphic with some part of E. The tangent space at a point of U is E, so we have $TU = U \times E$ as a coordinate neighbourhood for TM. To construct the cotangent bundle $P = T^*M$ we use the dual space of the tangent space in every point of U. This is E^* and $T^*U = U \times E^*$ is a neighbourhood for P. In the same way one constructs neighbourhoods for TP and T^*P. The vector Y_{p_q} with coordinates (x,α,e,β) can be represented by a curve $\ell(t) = (x+et, \alpha+t\beta)$ on P. The vector acts on functions on P in the following way

$$Y_{p_q}(f) = \lim_{t\to 0} \frac{f(x+et, \alpha+\beta t) - f(x,\alpha)}{t}. \tag{1.6.1}$$

After these preliminaries we define on P the canonical 1-form θ (as in (1.5.1) by:

$$\theta_{p_q}\left(Y_{p_q}\right) = \langle \pi_+ Y_{p_q} \mid p_q \rangle . \tag{1.6.2}$$

Expressed in natural coordinates we have (compare with (1.5.2))

$$\pi: P = T^*M \to M$$
$$(x,\alpha) \to x$$
$$\pi_*: TP \to TM$$
$$(x,\alpha,e,\beta) \to (x,e) .$$

The 1-form p_q has coordinates:

$$p_q = (x,\alpha) .$$

The right-hand side of (1.6.2) is therefore given by

$$\langle \pi_* Y_{p_q} \mid p_q \rangle = \langle (x,e) \mid (x,\alpha) \rangle = \langle e \mid \alpha \rangle . \tag{1.6.3}$$

Here is the last bracket the dual contraction between E and E^*. The left-hand side of (1.6.2) is (using the table):

$$\theta_{p_q}\left(Y_{p_q}\right) = \langle (e,\beta) \mid (\gamma,f) \rangle = \langle e \mid \gamma \rangle + \langle \beta \mid f \rangle \tag{1.6.4}$$

Equating (1.6.3) and (1.6.4) we obtain

$$\langle e \mid \alpha \rangle = \langle e \mid \gamma \rangle + \langle \beta \mid f \rangle .$$

This is solved by $f = 0$ and $\gamma = \alpha$.

Therefore the canonical 1-form is given by

$$\theta_{p_q} = (x,\alpha,\alpha, 0) . \qquad (1.6.5)$$

ompare this with (1.5.3).

Next we define the symplectic 2-form ω as in (1.5.4). Using the definition of exterior derivative we have

$$\omega_{p_q}(Y_1,Y_2) = d\theta_{p_q}(Y_1,Y_2) = (Y_1)_{p_q}\theta(Y_2) - (Y_2)_{p_q}\theta(Y_1)$$
$$-\theta_{p_q}([Y_1,Y_2]) \qquad (1.6.6)$$

with Y_1 and Y_2 vectors in $T_{p_q}P$.

To obtain the symplectic form in local coordinates we have to calculate the right-hand side of (1.6.6). To do this we have to extend the vectors Y_1 and Y_2 to vector fields in a neighbourhood of p_q . One can proof that the result, ω , is independent of the extension. Therefore we can choose a simple form for these vector fields, we will use vector fields that are constant in some local chart. So if

$$Y_1 = (x,\alpha, e_1,\beta_1)$$
$$Y_2 = (x,\alpha, e_2,\beta_2)$$

are the vectors in the point $p_q = (x,\alpha)$, then the vector fields in a point $\hat{p}_{\hat{q}} = (\hat{x},\hat{\alpha})$ in the neighbourhood have the same value:

$$Y_1(\hat{p}_{\hat{q}}) = (\hat{x},\hat{\alpha}, e_1,\beta_1)$$
$$Y_2(\hat{p}_{\hat{q}}) = (\hat{x},\hat{\alpha}, e_2,\beta_2).$$

In this case the commutator of $[Y_1,Y_2]$ vanishes and we are left with

$$\omega_{p_q}(Y_1,Y_2) = (Y_1)_{p_q}\theta(Y_2) - (Y_2)_{p_q}\theta(Y_1). \tag{1.6.7}$$

From (1.6.5) we obtain:

$$\theta_{p_q}(Y_2) = \langle e_2 \mid \alpha \rangle_{(x,\alpha)}. \tag{1.6.8}$$

The label (x,α) is attached to express the fact that (1.6.8) is a function on P. The action of Y_1 on this function is, using (1.6.1):

$$\begin{aligned}(Y_1)_{p_q}\theta(Y_2) &= \lim_{t\to 0}\frac{\theta(Y_2)(x+e_1t,\ \alpha+\beta_1 t) - \theta(Y_2)(x,\alpha)}{t} \\ &= \lim_{t\to 0}\frac{\langle e_2 \mid \alpha+\beta_1 t\rangle - \langle e_2 \mid \alpha\rangle}{t} = \langle e_2 \mid \beta_1\rangle .\end{aligned}$$

Similarly we have:

$$(Y_2)_{p_q}(\theta(Y_1)) = \langle e_1 \mid \beta_2 \rangle$$

and the result reads

$$\omega_{p_q}(Y_1,Y_2) = \langle e_2 \mid \beta_1\rangle - \langle e_1 \mid \beta_2 \rangle \tag{1.6.9}$$

with

$$p_q = (x,\alpha),\quad (Y_1)_{p_q} = (x,\alpha,\ e_1,\beta_1),\quad (Y_2)_{p_q} = (x,\alpha,\ e_2,\beta_2).$$

Having defined ω we can now proceed as in section 1.4 to define the Hamiltonian vector field X_H.

1.7. Hamiltonian Systems.

In our discussion until now we have considered phase spaces which are cotangent bundles of some configuration space M. For general purposes this is too restrictive, there are examples of interesting dynamical systems that do not have a configuration space but do have a phase space. An example is that of a classical particle with spin, studied by Souriau, [7], p.180.

This leads us to define a *Hamiltonian system* as a triple (P,ω,H). The phase space P is a Banach manifold, ω is a symplectic form, i.e. a closed 2-form on P which induces an isomorphism between TP and T^*P. The pair (P,ω) is called a symplectic manifold. H is the Hamiltonian.

The Hamiltonian vector field X_H is defined by

$$i_{X_H}\omega = -dH$$

and the physical trajectories of the system are integral curves of X_H. This is the general setting of dynamics we are going to use in the following chapters.

Chapter 2

PRESYMPLECTIC MANIFOLDS, CONSTRAINT MANIFOLDS AND REDUCED PHASE SPACE

2.1. Introduction.

In Chapter 1 we discussed the mapping Ω (equation (1.4.3)) from the tangent bundle of phase space, TP, to its cotangent bundle, T^*P, induced by a closed 2-form ω. The mapping Ω was supposed to be an isomorphism between the fibers T_uP and T_u^*P for all points u of P. In that case the equation

$$i(X_H)\omega = -dH \tag{2.1.1}$$

for the Hamiltonian vector field X_H has a unique solution.

There are, however, many interesting situations in which the 2-form ω (or the mapping Ω) is not so well-behaved. For instance, it may be the case that there is a point u of P and a vector Y_1 (unequal to zero) in T_uP such that

$$\omega(Y_1,Y_2) = 0 \tag{2.1.2}$$

for <u>all</u> Y_2 in T_uP. The mapping Ω associated to this ω is then <u>not</u> injective: the vectors X and $X+Y_1$ have the same image in T_u^*P. This means that equation (2.1.1) has no unique solution in u. In general Ω will then also be not surjective, and there may be no solution at all to (2.1.1) in u.

One distinguishes four classes of closed 2-forms, depending on the properties of the induced mapping Ω: ω is called

a) *degenerate* if Ω is neither surjective nor injective

b) *weakly degenerate* if Ω is surjective

c) *weakly nondegenerate* or *weakly symplectic* iſ Ω is injective

d) *symplectic* or *nondegenerate* if Ω is an isomorphism

For finite dimensional manifolds the cases b), c) and d) are identical (TP and T^*P have the same dimension), but for infinite dimensional manifolds there could be a difference. Case b) seems not to be of physical interest and we will not consider it in the rest of this paper. If ω belongs to class a) or c) one calls the pair (P,ω) a *presymplectic* manifold, if ω belongs to class d) one speaks of a *symplectic* manifold.

Using this terminology we can say that equation (2.1.1) does not have a solution on all of P if (P,ω) is presymplectic. There may, however, be a submanifold of P (a *constraint* manifold) on which (2.1.1) does have a solution. One can then try to use this submanifold as the phase space of the dynamical system.

A systematic algorithm to find such a submanifold on which solutions of (2.1.1) do exist was developed by Gotay, Nester and Hinds ([9], [11], [12], see also [10], [13]), who geometrized and generalized the earlier Dirac-Bergmann theory of constraint ([8], [28], see also [1]), which was formulated in the traditional coordinate dependent formalism of mechanics.

We sketch here the general idea of this algorithm, the precise mathematical formulation will be given in section 2.3.

We start with a presymplectic manifold (P,ω) and a Hamiltonian H. Since Ω is not surjective the image $\Omega(T_uP)$ of a fiber T_uP is in general a subspace of T^*_uP and the 1-form dH is in general not in the range of Ω. See fig. 2.

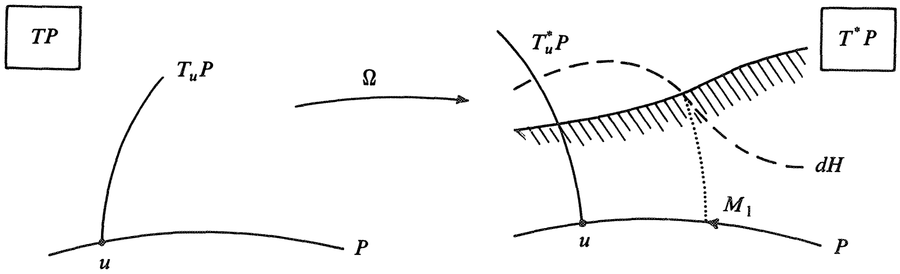


Fig. 2. The mapping Ω in the presymplectic case. The shaded region is the range of Ω, the dashed line is dH. The region of P where dH lies in the range of Ω is the constraint manifold M_1.

We now restrict ourselves to a subset M_1 where dH is in the range of Ω. We assume that M_1 is a manifold and we refer to it as a *constraint* manifold.

On M_1 we do have a solution X_H of the equations of motion (2.1.1), but X_H is not necessarily tangent to M_1: the vector $X_H(u)$ belongs to $T_uP\,|_{M_1}$ of which T_uM_1 is only a subspace. (See for a further discussion of this point Remark 1 following definition (2.2.1) and fig. 3). If X_H is not tangent to M_1 an integral curve of it will leave M_1. This means, interpreting the integral curves of X_H as the dynamical trajectories of our system, that the system will evolve to a point where (2.1.1) does not have a solution. Since this is unacceptable we restrict ourselves to a submanifold M_2 of M_1 where X_H is tangent to M_1. M_2 is the next constraint manifold. On M_2 we have solution X_H, tangent to M_1, but not necessarily to M_2 itself. So we must again restrict ourselves to another constraint manifold M_3 of M_2 where X_H is tangent to M_2. On M_3 we can have the same problem (X_H not tangent to M_3) and in this way we are forced to construct a chain of submanifolds which hopefully ends at a submanifold M_k such that X_H is tangent to M_k in <u>all</u> points of M_k. It may happen that this final constraint manifold is empty. In that case we conclude that there is no way to define consistent dynamics on the original presymplectic manifold (P,ω) with the given Hamiltonian.

In most cases though the final constraint manifold will be nontrivial and in the next section we will develop the technical tools needed to construct explicitly the chain of constraint manifolds $M_1 M_2, \ldots, M_k$.

2.2. Technicalities [9].

In this section (P,ω) is a presymplectic manifold, M is a submanifold of P and

$$j: M \to P$$

is the embedding of M in P. We will often identify M and its image in P, but it is good to keep the difference between the two in mind.

The mapping Ω

$$\Omega: TP \to T^*P$$

is defined by

$$\Omega(X) = i(X)\omega \tag{2.2.1}$$

and Ω is assumed to be *closed*, i.e. Ω maps closed sets in closed sets.

DEFINITION 2.2.1

$$\underline{TM} = \{Y \in TP|_M \mid Y = j_* X,\ X \in TM\}$$

REMARK 1

$\underline{TM}$ is the embedding of TM into $TP|_M$, but it is not equal to $TP|_M$. Take for example $P = \mathbb{R}^3$, $M = S^2$ and $j: M \to P$ the standard embedding of the 2-sphere in $\mathbb{R}^3$. A vector in the radial direction is then an element of $TP|_M$ but not of $\underline{TM}$ (see fig. 3).

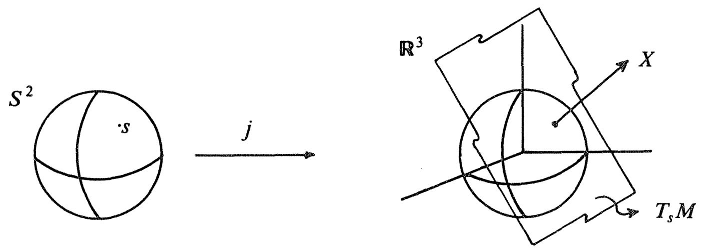


Fig. 3. The difference between $\underline{TM}$ and $TP|_M$:
X belongs to $TP|_M$ but not $\underline{TM}$.

REMARK 2.

$\underline{TM}$ is closed because j induces a homeomorphism between TM and $\underline{TM}$.

DEFINITION 2.2.2.

$$TM^{\perp} = \{X \in TP|_M \mid \omega(X,Y) = 0,\ \forall Y \in \underline{TM}\}$$

$TM^{\perp}$ is called the *symplectic complement* of TM with respect to ω .

We now need some facts about Banach spaces. Let S be a subspace of the Banach space E .

DEFINITION 2.2.3.

The *annihilator* of S is the subspace of the dual of E defined by

$$S^{\vdash} = \{V \in E^* \mid V(s) = 0,\ \forall s \in S\} .$$

Let L be a subspace of E^* .

DEFINITION 2.2.4.

The *annihilator* of L is the subspace of E defined by

$$L^{\vdash} = \{e \in E \mid \ell(e) = 0,\ \forall \ell \in L\}$$

PROPOSITION 2.2.5.

$$(S^{\vdash})^{\vdash} = \bar{S} \, .$$

PROOF.

1) Let $v \in S^{\vdash}$, then $v(s) = 0$ for all s in S and by continuity $v(t) = 0$ for all t in the closure of S ; accordingly $\bar{S}$ annihilates any $v \in s^{\vdash}$ and

$$\bar{S} \subset (S^{\vdash})^{\vdash} .$$

2) Suppose there is an element w of $(S^{\vdash})^{\vdash}$ not in $\bar{S}$, then by the Hahn-Banach theorem (see for example [27], theorem 12.3) there exists a continuous linear functional β such that

$$\beta(w) \neq 0 \quad \text{and} \quad \beta(t) = 0 \, , \quad \text{for all } t \text{ in } \bar{S} \, .$$

From $\beta(t) = 0$ follows that β is an element of $S^{\vdash}$ and therefore $\beta(w) = 0$. This is a contradiction and we conclude that an element w in $(S^{\vdash})$ but not in $\bar{S}$ does not exist:

$$(S^{\vdash})^{\vdash} \subset \bar{S} \, . \qquad \blacksquare$$

PROPOSITION 2.2.6.

Let E *be reflexive, i.e.* $E^{**} = E$ *and* L *a subspace of* E^{*} ; *then*

$$(L^{\vdash})^{\vdash} = \bar{L} \, .$$

The proof of this proposition is analogous to the one of the previous proposition.

We can extend the notion of annihilator and symplectic complement to fiber bundles, the fibers of which are Banach spaces, by calculating the annihilator or symplectic complement in each fiber separately.

We now formulate the basic theorem of the constraint algorithm, which gives a characterization of the range of the mapping Ω induced by the closed 2-form ω.

THEOREM 2.2.7.

$$\Omega(\underline{TM}) = (TM^{\perp})^{\vdash}.$$

PROOF.

We start by noting that

$$\Omega(\underline{TM})^{\vdash} = TM^{\perp}, \tag{2.2.2}$$

since by definition of the annihilator (Definition 2.2.3)

$$\Omega(\underline{TM})^{\vdash} = \{Y \in TP|_M \mid \langle Y \mid \Omega(X) \rangle = 0, \ \forall X \in \underline{TM}\}$$

and by definition of Ω (equation (2.2.1))

$$\langle Y \mid \Omega(X) \rangle = \omega(X,Y).$$

So we obtain

$$\Omega(\underline{TM})^{\vdash} = \{Y \in TP|_M \mid \omega(X,Y) = 0, \ \forall X \in \underline{TM}\}$$

but this is the definition of $TM^{\perp}$ (cf. Definition 2.2.2). This proofs equation (2.2.2).

We then apply the annihilator to both sides of equation (2.2.2) to find, using Proposition 2.2.6 :

$$\overline{\Omega(\underline{TM})} = (\Omega(\underline{TM})^{\vdash})^{\vdash} = (TM^{\perp})^{\vdash}.$$

Now Ω was supposed to be a closed mapping and $\underline{TM}$ is closed, so finally we get:

$$\Omega(TM) = (TM^{\perp})^{\vdash} . \quad \blacksquare$$

An immediate consequence of this theorem is that a form α is in the range of $\Omega|_M$ if and only if

$$\langle TM^{\perp} \mid \alpha \rangle = 0 .$$

We will use this fact repeatedly to calculate the chain of constraint manifolds.

2.3. The constraint algorithm.

Let (P,ω) be a presymplectic manifold and H a Hamiltonian function:

$$H: P \rightarrow \mathbb{R} .$$

Our aim is to find a Hamiltonian vector field X_H on P, i.e. a solution of

$$i(X_H)\omega = -dH . \qquad (2.3.1)$$

This means that dH must be in the range of Ω,

$$dH \in \Omega(TP) . \qquad (2.3.2)$$

Using the fundamental Theorem 2.2.7 this is equivalent to

$$dH \in (TP^{\perp})^{\vdash}$$

or

$$\langle TP^{\perp} \mid dH \rangle = 0 . \qquad (2.3.3)$$

As Ω is not surjective for presymplectic ω equation (2.3.3) is not true for all points of P and we must restrict ourselves to points where it is true. This gives the first constraint manifold M_1 :

$$M_1 = \{p \in P \mid \langle TP^{\perp} \mid dH \rangle(p) = 0\} . \tag{2.3.4}$$

On M_1 we have at least one solution X_H of (2.3.1) but this solution need not be tangent to M_1, as we argued in section 2.1. For a solution to be tangent to M_1 we must have

$$dH \in \Omega(\underline{TM}_1) .$$

Using again the fundamental theorem this leads to the condition

$$\langle TM_1^{\perp} \mid dH \rangle = 0 .$$

This need not be true on the whole of M_1 and we have to make a further restriction to a submanifold M_2 of M_1 :

$$M_2 = \{m \in M_1 \mid \langle TM_1^{\perp} \mid dH \rangle(m) = 0\} .$$

In this way we construct a chain of submanifolds

$$M_\ell = \{m \in M_{\ell-1} \mid \langle TM_{\ell-1}^{\perp} \mid dH \rangle(m) = 0\} .$$

If we end up with a manifold M_k with the property

$$\langle TM_k^{\perp} \mid dH \rangle = 0$$

for all of its points the chain stops. By construction there will be a solution of (2.3.1) on M_k, for

$$dH \in \Omega(\underline{TM}_k)$$

and this solution will be tangent to M_k.

M_k is called the final constraint manifold.

2.4. Gauge freedom.

In this section we consider the equations of motion on the final constraint manifold M_k :

$$i(X_H)\ \omega/M_k = -dH/M_k\ . \tag{2.4.1}$$

On M_k this equation has a solution X_H but if Ω is not injective the solution need not be unique. Any vector $Z \in \ker \cap \underline{TM}_k$ may be added to X_H to give a new solution

$$\tilde{X}_H = X_H + Z \tag{2.4.2}$$

which is again tangent to M_k .

One refers to this arbitrariness as *gauge freedom* and X_H and $X_H + Z$ are called *gauge equivalent vectors*. Using the physical interpretation of integral curves of solutions of (2.4.1) as dynamical trajectories, gauge freedom means that an initial condition, a point $m(o)$ of M_k , will evolve to different points $m_{X_H}(t)$, depending on the choice for the solution of (2.4.1).

We now make the crucial assumption that this freedom of choice has no physical significance. The solutions X_H and $X_H + Z$ describe the same physics. This leads to the following

DEFINITION 2.4.1.

Points on M_k are *physically equivalent* (or *gauge equivalent*) if they can be reached from the same initial condition by integral curves of solutions of equation (2.4.1) in the same lapse of "time". By "time" we mean the variable t parametrizing the integral curves.

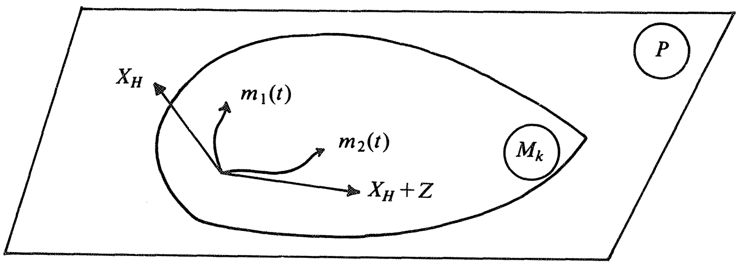


Fig. 4. Gauge freedom.
X_H and $X_H + Z$ are gauge equivalent vector field; $m_1(t)$ and $m_2(t)$ are integral curves of X_H and $X_H + Z$. For each t the points $m_1(t)$ and $m_2(t)$ are equivalent.

COROLLARY 2.4.2.

Points that can be reached in the same time from equivalent points are equivalent.

We now define *gauge vector fields.*

DEFINITION 2.4.3.

A gauge vector field is a vector field whose integral curves consist of physically equivalent points.

The collection

$$G_1 = \ker \Omega \cap \underline{TM}_k$$

considered above consists of gauge vector fields, but there are (many) more. This is shown by the following proposition.

PROPOSITION 2.4.4.

Let X *be a solution of equation* (2.4.1) *and* Z *a gauge vector field. Then their commutator* [X,Z] *is also a gauge vector field.*

PROOF.

Let ψ_X^t and ψ_Z^t be the flows (local diffeomorphisms) associated with the vector fields X and Z. Consider a point m_o in M_k. Applying the flow ψ_Z^t to m_o we get a gauge equivalent point $\psi_Z^t(m_o)$. The evolution of these two points generated by X is

$$\psi_X^t : m_o \longmapsto m_1 = \psi_X^t(m_o)$$

$$\psi_Z^t(m_o) \longmapsto m_2 = \psi_X^t(\psi_Z^t(m_o))$$

The points m_1 and m_2 are equivalent (use the Corollary 2.4.2). The point m_2 is equivalent to $\psi^{-t}_Z(m_2)$ and accordingly m_1 and $\psi^{-t}_Z(m_2)$ are equivalent. Applying finally ψ^{-t}_X to m_1 and $\psi^{-t}_Z(m_2)$:

$$\psi^{-t}_X : m_1 \longmapsto \psi^{-t}_X(m_1) = \psi^{-t}_X(\psi_X^t(m_o)) = m_o$$

$$\psi^{-t}_Z(m_2) \longmapsto m_3(t) = \psi^{-t}_X\psi^{-t}_Z(m_2) = \psi^{-t}_X\psi^{-t}_Z\psi^t_X\psi^t_Z(m_o)$$

we obtain the gauge equivalent of m_o and $m_3(t)$ for all t.
The curve $m_3(t)$ consists of points equivalent to m_o and hence of points equivalent to each other. This means that the tangent vector to $m_3(t)$ for $t = 0$ is a gauge vector. The tangent vector to $m_3(t)$ is the commutator $[X,Z]$ (see for example Theorem 4 of section 1.4 of [29]), which is therefore a gauge vector field. ∎

A similar argument shows that if Z_1 and Z_2 are gauge vector fields then also the commutator $[Z_1,Z_2]$ is a gauge vector field.

Starting from $G_1 = \ker\Omega \cap \underline{TM}_k$ we can construct a series of sets $G_2, G_3 \ldots G_\ell$ of gauge vector fields by defining:

$$G_\ell = G_{\ell-1} + [X,G_{\ell-1}] + [G_{\ell-1},G_{\ell-1}] \qquad \ell = 2,3,\ldots$$

with X a solution of 2.4.1 tangent to M_k. Obviously one has

$$G_{\ell-1} \subset G_\ell ,$$

so by this construction we enlarge in every step the collection of gauge vector fields until for some ℓ_f

$$G_{\ell_f - 1} = G_{\ell_f} .$$

Then this proces stops and we have

$$G_k = G_{\ell_f}$$

for all $k > \ell \cdot G_{\ell_f}$ is then the complete set of gauge vectors.

One can proof that for all ℓ

$$G_\ell \subset TM_k^\perp \cap \underline{TM}_k .$$

Because the proof is somewhat technical we give it in Appendix A.

To proceed we make the following assumption:

<u>Assumption 2.4.5.</u>

Define N by

$$N = TM_k^\perp \cap \underline{TM}_k$$

then we assume

$$\lim_{\ell \to \infty} G_\ell = N . \qquad (2.4.3)$$

In most cases equation (2.4.3) is true but one can construct some counter examples (see [11], [12]).

One should realize that a vector $X_H + X$ (with X_H a solution of (2.4.1) and X an element of $TM_k^\perp$) is only another solution of the equations of motion tangent to M_k if both

a) $X \in \underline{TM_k}$, so X must belong to N, and

b) we restrict the vector space TP/M_k on which the forms dH/M_k and $i(X_H+X)\omega/M_k$ act to its subspace $\underline{TM_k}$, for only then we have

$$\omega(X_H+X,Y) = -dH(Y)$$
$$= \omega(X_H,Y) = -dH(Y)$$

because $X \in TM_k$ satisfies

$$\omega(X,Y) = 0$$

only for Y in $\underline{TM_k}$ and not for arbitrary Y in TP/M_k. The restriction of TP/M_k to $\underline{TM_k}$ is equivalent to pulling back the equations of motion to M_k and so the assumption 2.4.5 is tantamount to this pulling back to M_k. (cf. [11]).

The assumption 2.4.5 enables us to characterize the final constraint manifold M_k by its gauge freedom, that is by N, or more geometrically by the way TM_k intersects $\underline{TM_k}$. One introduces the following terminology ([30], [2]):

M_k is called

a) *Isotropic* if $\underline{TM_k} \subset TM_k^{\perp}$

b) *Coisotropic* (or *first class*) if $TM_k^{\perp} \subset \underline{TM_k}$

c) *Second class* (or *weakly symplectic*) if $\underline{TM_k} \cap TM_k^{\perp} = \{0\}$

d) *Mixed* in all other cases.

If the manifold M_k is both isotropic and coisotropic (i.e. $\underline{TM_k} = TM_k^{\perp}$) it is called a *Lagrangian* submanifold. Lagrangian submanifolds play an important role in the theory of canonical transformations and in geometric quantization ([2]), but are not very interesting in the context of constraint manifolds, see point a) below. The name weakly symplectic for case c) will be explained in section 2.5.

From the point of view of gauge freedom the above characterization of M_k has the following meaning.

a) An isotropic constraint manifold has as its gauge vectors all tangent vectors. This means that all points are gauge equivalent and that there is no dynamics on M_k.

b) A first class manifold has maximal gauge freedom: all potential gauge vectors (i.e. vectors in $TM_k^{\perp}$) are in fact gauge vectors (i.e. belong to $\underline{TM_k}$).

c) In a second class manifold there is no gauge freedom, all points are physically inequivalent.

As for the case of a mixed constraint manifold we refer to a theorem by Sniatycki [13], which says that every mixed constraint manifold M_k of P is a first class submanifold with respect to some submanifold $\tilde{P}$ of P, with $\tilde{P}$ itself a second class manifold of P.

These considerations show that for our purposes (finding physical phase spaces from presymplectic systems) we can restrict our attention to first and second class submanifolds. These cases will be discussed in the next two sections.

2.5. Second class constraint manifolds.

The final constraint manifold M_k of a presymplectic system (P,ω,H) is called second class (or weakly symplectic) if

$$\underline{TM_k} \cap TM_k^{\perp} = \{0\} . \qquad (2.5.1)$$

For such manifolds there are no gauge vector fields and the equation

$$i(X_H)\omega\Big|_{M_k} = -dH\Big|_{M_k} \qquad (2.5.2)$$

has a unique solution tangent to M_k.

The name "weakly symplectic manifold" can be understood from the following theorem.

THEOREM 2.5.1.

The pullback $\omega_k = j^*\omega$ *of the presymplectic form* ω *to the final constraint manifold* M_k *by the submanifold map* j *is weakly symplectic, i.e. the associated mapping*

$$\Omega_k : TM_k \to T^*M_k$$

is injective.

PROOF.

Suppose $\omega_k(X,Y) = 0$ for some $X(\neq 0)$ and all Y in $\underline{TM_k}$. Then $\omega(j_*X,j_*Y) = 0$ for all $j_*Y \in \underline{TM_k}$. Consequently $j_*X \in TM_k^{\perp}$. Of course, j_*X is also an element of $\underline{TM_k}$ and we conclude from the definition of a weakly symplectic manifold that $j_*X = 0$. Since j_* is injective X must be zero, hence $j^*\omega$ is weakly symplectic (and Ω_k is injective). ∎

We want to describe the dynamics entirely in terms of objects defined on M_k. Therefore we pull the equations of motion (2.5.2) back to M_k by the submanifold map to obtain:

$$j^*[i(X_H)\omega\big|_{M_k}] = j^*[-(dH)\big|_{M_k}]. \tag{2.5.3}$$

We use for the right-hand side

$$j^*[(dH)\big|_{M_k}] = d(j^*H)\big|_{M_k} = d(H\big|_{M_k}) \tag{2.5.4}$$

Equation (2.5.4) is not entirely trivial, because we take the pullback of a restricted form $dH\big|_{M_k}$, and restricted forms are somewhat peculiar. For instance, the exterior derivative of the restriction of an exact form is not necessarily zero:

$$d\left((dH)\Big|_{M_k} \right) \neq 0$$

in general. For a discussion of $(dH)\big|_{M_k}$ see Appendix A.

The left-hand side of equation (2.5.3) is a form in T^*M_k acting on a vector Y in TM_k in the following way

$$j^*[i(X_H)\omega\Big|_{M_k}](Y) = \omega(X_H, J_*Y) \tag{2.5.5}$$

We know that X_H lies in $\underline{TM}_k$ (by construction of the final constraint manifold M_k). This means that there is a vector $\tilde{X}_H$ in TM_k, with

$$X_H = j_*\tilde{X}_H .$$

So we find for the right-hand side of equation (2.5.5):

$$\begin{aligned} \omega(X_H, j_*Y) &= \omega(j_*\tilde{X}_H, j_*Y) = j^*\omega(\tilde{X}_H, Y) \\ &= [i(\tilde{X}_H)j^*\omega](Y) \end{aligned} \tag{2.5.6}$$

Introducing the notation

$$\omega_k = j^*\omega$$
$$H_k = j^*H = H\Big|_{M_k}$$

and using equations (2.5.4), (2.5.5) and (2.5.6) we obtain for the "pulled back" equations of motion.

$$i(\tilde{X}_H)\omega_k = -dH_k \tag{2.5.7}$$

Here we have expressed the equations of motion in terms of intrinsically defined objects on M_k. In this derivation of (2.5.7) we have nowhere used the fact that M_k is second class, so (2.5.7) is true for arbitrary final constraint manifolds. However, for a second class constraint manifold (2.5.7) has a unique solution.

The conclusion of this section is that for a presymplectic system with a second class final constraint manifold we can forget that we started with a presymplectic triple (P,ω,H) and use the weakly symplectic system (M_k,ω_k,H_k) as the description of the system. We can say that we have eliminated unphysical degrees of freedom by going from P to the "true" phase space M_k. In the next section we will show that in the case of a first class constraint manifold we have to eliminate even more degrees of freedom to get rid of the gauge freedom of the final constraint manifold.

In the rest of this paper we will not discuss second class constraint manifolds any further. We would like to remark that there are very interesting physical systems described by second class manifolds. An example is the Korteweg-de Vries equation (see McFarlane [14]).

2.6. First class constraint manifolds.

The final constraint manifold M_k of a presymplectic system (P,ω,H) is called first class if

$$TM_k^{\perp} \subset \underline{TM}_k \ . \tag{2.6.1}$$

For such manifolds the solutions of the equations of motion are not unique. We can add to any solution X_H tangent to M_k (such solutions do exist by the construction of M_k) an element Z of $TM_k^{\perp}$ to get another vector field which is according to (2.6.1) again tangent to M_k. We argued (or rather postulated) in section 2.4 that X_H and $X_H + Z$ describe the same physics, and that points connected by the integral curves of gauge vector fields $Z \in TM_k^{\perp} \cap \underline{TM}_k$ are physically equivalent.

To eliminate this gauge freedom we want to identify gauge equivalent points of M_k. We therefore have to find the set L_m of all points that are equivalent to a given point m of M_k. One constructs L_m by taking the union of all integral curves of vectors Z going through m.

According to the Frobenius theorem (see for instance Theorem 2.2.26 of [2]) the set L_m is a smooth manifold (a *maximal integral manifold* of $TM_k^\perp$) if $TM_k^\perp$ is involutive i.e. if for any pair Z_1, Z_2 in $TM_k^\perp$ their commutator $[Z_1,Z_2]$ is also in $TM_k^\perp$. Fortunately, we have the following theorem:

THEOREM 2.6.1.

$TM_k^\perp$ *is involutive.*

PROOF.

Let Z_1 and $Z_2 \in TM_k^\perp$, this means that

$$i(Z_i)\omega(j_*X) = 0 , \qquad i = 1,2 , \qquad \forall X \in TM_k .$$

We have to show

$$i([Z_1,Z_2])\omega(j_*X) = 0 .$$

Using the indentities A.7 and A.8 from Appendix A we calculate

$$i([Z_1,Z_2])\omega(j_*X) = \left[L_{Z_1}\, i(Z_2) - i(Z_2)\, L_{Z_1}\right]\omega(j_*X) =$$

$$[di(Z_1)i(Z_2) + i(Z_1)di(Z_2) - i(Z_2)i(Z_1)d - i(Z_2)di(Z_1)]\omega(j_*X) = 0 .$$

All terms are separately zero, the first, second and fourth because Z_1 and Z_2 belong to $TM_k^\perp$ and the third because ω is closed. ∎

Because of this theorem and the theorem of Frobenius there goes through every point m of M_k a maximal integral manifold L_m of the collection of gauge vector $TM_k^\perp$. One says that M_k is "foliated" by the "leaves" L_m of gauge equivalent points (see [2], [3] and fig. 5).

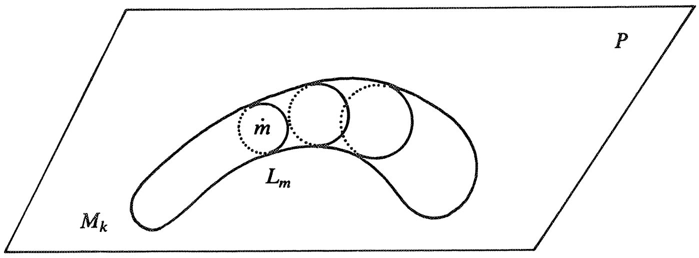


Fig. 5. The foliation of M_k by leaves of gauge equivalent points.

We can now form the quotient space $R = M_{k/\sim}$ of M_k by the following equivalence relation: two points m_1 and m_2 of M_k are equivalent if they lie on the same leaf L, or in other words, if they are gauge equivalent. R is the space of leaves L_m of M_k.

We have a canonical projection

$$\pi: M_k \to R$$

which assigns to a point of M_k the leaf to which it belongs. We assume R to be a smooth manifold and π to be a smooth mapping. R is called the *reduced phase space* of the presymplectic system (P,ω,H).

We can express the dynamics entirely in terms of objects defined on R. We first introduce on R a 2-form ω_R in the following way. Let Y_1 and Y_2 be two vectors tangent to a point ℓ of R. A point ℓ of R corresponds to a leaf L in M_k. Take any point m on L and two tangent vectors $\tilde{Y}_1$ and $\tilde{Y}_2$ in T_mM_k that project on Y_1 and Y_2:

$$Y_1 = \pi_*\tilde{Y}_1 , \qquad Y_2 = \pi_*\tilde{Y}_2 .$$

We define ω_R by

$$\omega_R(Y_1,Y_2) = \omega_k(\tilde{Y}_1,\tilde{Y}_2)$$

with $\omega_k = j^*\omega$

the pullback of ω onto M_k.

THEOREM 2.6.2.

ω_R *is well-defined.*

PROOF.

We have to show that the definition is independent of

a) the choice of point m on L,

b) the choice of tangent vectors $\tilde{Y}_1$ and $\tilde{Y}_2$.

Concerning a) we note that ω_k is constant along a leaf: take any tangent vector field to a leaf, i.e. any element Z of $TM_k^\perp$, then

$$L_Z\omega_k = i(Z)d\omega_k + di(Z)\omega_k = 0$$

because ω_k is closed and $Z \in TM_k^\perp$.

To proof b) we remark that if we choose two other vectors Y_1 and Y_2 that project on Y_1 and Y_2:

$$Y_1 = \pi_*\hat{Y}_1, \qquad Y_2 = \pi_*\hat{Y}_2$$

then

$$\hat{Y}_1 = \tilde{Y}_1 + Z_1, \qquad \hat{Y}_2 = \tilde{Y} + Z_2$$

with Z_1 and Z_2 elements of $TM_k^\perp$. Using this we obtain

$$\omega_k(\hat{Y}_1,\hat{Y}_2) = \omega_k(\tilde{Y}_1+Z_1, \tilde{Y}_2+Z_2) = \omega_k(\tilde{Y}_1,\tilde{Y}_2). \blacksquare$$

THEOREM 2.6.3.

ω_R *is weakly symplectic.*

PROOF.

Suppose

$$\omega_R(Y_1,Y_2) = 0$$

for all $Y_2 \in TR$, then

$$\omega_k(\tilde{Y}_1,\tilde{Y}_2) = 0$$

for all $\tilde{Y}_2 \in TM_k$ or consequently $\tilde{Y}_1 \in TM_k^{\perp}$. But then the projection of $\tilde{Y}_1$ is zero:

$$Y_1 = \pi_*\tilde{Y}_1 = 0. \quad ∎$$

THEOREM.

The Hamiltonian H_k *is constant along leaves of the foliation. (One can also express this by saying that* H_k *is gauge invariant).*

PROOF.

Let X be tangent to a leaf, i.e. $X \in TM_k^{\perp}$. Then, using the definition of X_H we calculate

$$X(H_k) = dH_k(X) = -\omega_k(X_H,X) = 0$$

(because $X \in TM_k^{\perp}$). H_k is constant along leaves. ∎

On R we have in this way constructed a weakly symplectic for ω_R and a well-defined Hamiltonian H_R. $(H_R(L_m) := H_k(m))$. We can formulate the equations of motion on R:

$$i(X_{H_R})\omega = -dH_R$$

The unique solution X_{H_R} of this equation is the projection of all solutions of the equations of motion on M_k . This is because two solutions of M_k differ precisely by an element of $TM_k^{\perp}$, which is "projected out" by going to R .

With the reduced phase space R we have found the true phase space of our original presymplectic system (P,ω,H) . In the next chapter we will determine the character of R for the Yang-Mills system.

Chapter 3

THE YANG-MILLS SYSTEM

3.1. Introduction.

We will calculate in this chapter the final constraint manifold of the Yang-Mills system using the techniques developed in the last chapters.

First we recall some definitions and results on the fiber bundle formulation of gauge theories. For more details see Eguchi et. al. [17], Bleeker [16] and Daniel and Viallet [18].

Then, in section 3.3, we show that the Yang-Mills equations can be formulated as a presymplectic Hamiltonian system. In section 3.4 we apply the constraint algorithm to calculate the final constraint manifold. In section 3.5 we discuss the gauge freedom and the equations of motion on the final constraint manifold and in section 3.6 we calculate the reduced phase space and conclude this report with some general remarks.

3.2. Fiber bundles.

3.2.1. Principle fiber bubdles.

A *principle fiber bundle* $P(M,G)$ is a manifold P on which the Lie group G acts freely and differentially from the right: there is a mapping

$$\Phi\colon P\times G \to P$$

such that

1) Φ is smooth
2) $\Phi(\Phi(p,g_1),g_2) = \Phi(p,g_1g_2)$, $\forall p \in P$, $\forall g_1,g_2 \in G$
and
$\Phi(p,e) = p$, $\forall p \in P$, e is the identity element of G
3) if $\Phi(p,g) = p$ for any $p \in P$, then $g = e$.

One also uses the notation

$$\Phi(p,g) = pg = \Phi_g(p) .$$

This action of G on P defines an equivalence relation on P : two points p_1 and p_2 of P are equivalent if there exists a $g \in G$ such that

$$p_1 = p_2 g .$$

The quotient space P/G is M . One further assume that the projection

$$\pi : P \rightarrow P/G = M$$

which assigns to an element of P its equivalence class in M is such that there exists on M an open covering $\{U_\alpha\}$ such that

$$\pi^{-1}(U_\alpha) \simeq U_\alpha \times G .$$

Thus P is locally trivial, i.e. the product of M and G and we can locally (on $\pi^{-1}(U_\alpha)$) introduce coordinates in P , such that

$$p = (x,g), \text{ with } p \in \pi^{-1}(U_\alpha), \; x = \pi(p) \in U_\alpha, \; g \in G .$$

The action of G takes in these coordinates the simple form:

$$pg_1 = (x,g)g_1 = (x,gg_1) .$$

One calls P the *total space* of the principle fiber bundle, M is the *base-space*, G the *structure group* and $\pi^{-1}(m)$, with $m \in M$, the *fiber* over m . Tangent vectors X_p in T_pP are called *vertical* if $\pi_* X_p = 0$: they are tangent to the fiber over $\pi(p)$.

3.2.2. Associated fiber bundles.

Let P(M,G) be a principle fiber bundle as defined above. We construct for every linear representation

$$\text{Rep}: G \to GL(F)$$

of G on some vector space F a new manifold E, that has locally the product structure $M \times F$ in the following way: Define on $P \times F$ the action of G by

$$(p,f) \to (p,f)g = (pg, \text{Rep}(g^{-1})f), \quad p \in P, \quad f \in F, \quad g \in G .$$

Choosing local coordinates on P we can write

$$(p,f) = (x,g,f)$$

and the action on G takes the form

$$(p,f)g_1 = (pg_1, \text{Rep}(g_1^{-1})f) = (x,gg_1, \text{Rep}(g_1^{-1})f) .$$

The action of G on $P \times F$ defines an equivalence relation in the usual way:

$$(p_1,f_1) \sim (p_2,f_2) \Longleftrightarrow \exists_g (p_1 g, \text{Rep}(g^{-1})f_1) = (p_2,f_2) .$$

The quotient space $P \times F/G$ is called E.

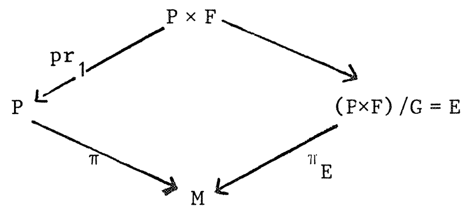
Using the local coordinates we can always choose a representative of an equivalence class to be of the form

$$(x,e,f)$$

by making a suitable transformation. So we find that locally E can be coordinatizaed by (x,f) and E is locally a product $M \times F$. More abstractly: there exists a projection

$$\pi_E: E \to M$$

(defined by the commutativity of the following diagram:



pr_1 is the projection on the first coordinate.) such that

$$\pi_E^{-1}(U_\alpha) \cong U_\alpha \times F$$

with $\{U_\alpha\}$ the same open covering of M used in the definition of P(M,G).

E is called the vector bundle associated to P(M,G) and the representation Rep of G on F .

3.2.3. Sections.

A mapping

$$s: M \to P(M,G)$$

is called a *section* of the principle fiber bundle P iff

$$\pi \circ s = id_M .$$

The collection of all sections is denoted by Γ(P) . One similarly defines sections of an associated fiber bundle E . One denotes the collection of sections of E by Γ(E) .

Important is the following fact [18]: there is a bijective correspondence between Γ(E) and the collection of functions

$$f: P \to F$$

with

$$f(pg) = Rep(g^{-1})f(p) \tag{3.2.3.1}$$

To see this, introduce coordinates in P and E. A section s_E of E has the form

$$x \longmapsto s_E(x) = (x, \tilde{s}_E(x)) = [(x,e,\tilde{s}_E(x))]$$
$$= [(x,g,\mathrm{Rep}(g^{-1})\,\tilde{s}_E(x))]$$

with $\tilde{s}_E(x) \in F$ and square brackets denoting the equivalence classes in $P \times F$. We have for every point $p = (x,g)$ in P a value $\mathrm{Rep}(g^{-1})\,\tilde{s}_E(x)$ in F. This defines the function (3.2.3.1). If on the other hand we have a function (3.2.3.1) then we can define a section of E by

$$x \longrightarrow [(x,g,f(x,g))] = [(x,e,\mathrm{Rep}(g)\,\mathrm{Rep}(g^{-1})f(x,e))]$$
$$= [(x,e,f(x,e))] = (x,f(x,e)) .$$

One can generalize this result in the following way: Let $\Lambda^k(M,\Gamma(E))$ be the collection of k-forms on M with values in the sections of E, and let $\bar{\Lambda}^k(P,F)$ be the set of k-forms on P with values in F such that for all $\alpha \in \bar{\Lambda}^k(P,F)$

1) $\Phi_g^* \alpha = \mathrm{Rep}(g^{-1})$,
2) α is zero on vertical vectors (vertical vectors were defined in the last paragraph of section 3.2.1).

Then the spaces $\Lambda^k(M, (E))$ and $\bar{\Lambda}^k(P,F)$ are isomorphic. The proof of these correspondences is analogous to the proof in the case of functions. One calls a form α transforming as in 1) a form of type Rep and a form α satisfying 2) is called horizontal.

3.2.4. Connections on a principle fiber bundle.

We defined in section 3.2.1 a tangent factor X_p in T_pP to be vertical if it is tangent to the fiber $\pi^{-1}(\pi(P))$. One can show that every vertical vector X_p is of the form

$$X_p = \frac{d}{dt}(p \cdot \exp t\,\Sigma)_{t=0}$$

for some Σ in the Lie algebra $\ell(G)$ of G .

To define the notion of *horizontal vectors* we introduce a so-called *connection form* A on P . A is a Lie algebra valued 1-form with the following properties

1) $\Phi_g^* A = ad(g^{-1})A$,

   ad is the adjoint representation of G , acting on the Lie algebra $\ell(G)$.

2) $A(\Sigma^*) = \Sigma$.

   Σ^* is the fundamental vertical vector field on P , induced by Σ in $\ell(G)$:

$$\Sigma^*_p = \frac{d}{dt}(p \cdot \exp(t\,\Sigma))_{t=0} .$$

We define a vector X to be horizontal (with respect to the connection form A) if

$$A(X) = 0 .$$

Property 1) says that a connection form is of type ad and property 2) tells us that A is <u>not</u> horizontal. However, in the next section we will show the relation between connection forms and the set $\bar{\Lambda}^1(P,\ell(G))$ of horizontal forms of type ad .

3.2.5. The space of all connections of a principle fiber bundle.

In general there are in a principle fiber bundle many ways to define horizontal vectors, i.e. there are many possible choices of connection form A .

We are interested in the space of all connections $C(P)$. This is an affine space modelled on the vector space $\bar{\Lambda}^1(P,\ell(G))$: consider two connection forms A_1 and A_2. Their difference

$$T = A_1 - A_2$$

is an element of $\bar{\Lambda}^1(P,\ell(G))$ for

1) $\Phi_g^* T = \Phi_g^* A_1 - \Phi_g^* A_2 = ad(g^{-1})(A_1 - A_2) = ad(g^{-1})T$

2) $T(\Sigma^*) = A_1(\Sigma^*) - A_2(\Sigma^*) = \Sigma - \Sigma = 0$
   for all vertical vectors Σ^*.

On the other hand, if A is a connection and T an element of $\bar{\Lambda}^1(P, (G))$ then $T + A$ is also a connection, as one easily checks.

$C(P)$ considered as a manifold has as its tangent space $T_A C(P)$ the model space $\bar{\Lambda}^1(P,\ell(G))$: the curves through a point A of $C(P)$ can be represented by

$$A_T(t) = A + tT$$

with

$$T \in \bar{\Lambda}^1(P,\ell(G)) .$$

3.2.6. Hodge star operator and metrics on sections.

In this section we collect some elementary facts of Riemannian geometry and use these to define a so-called Hodge star operator on $\Lambda^k(M)$, the space of k-forms on M. We extend this notion of the Hodge star operator to $\Lambda^k(M,\Gamma(E))$ the space of forms with values in the space of sections of some associated fiber bundle E with typical fiber F. Using the isomorphism between $\Lambda^k(M,\Gamma(E))$ and $\bar{\Lambda}^k(P,F)$ discussed in section 3.2.3 the Hodge star operator can be defined on $\bar{\Lambda}^k(P,F)$. Finally, the star operator is used to define metrics on the spaces

$$\Lambda^k(M), \quad \Lambda^k(M,\Gamma(E)) \quad \text{and} \quad \bar{\Lambda}^k(P,F) .$$

A. Riemann geometry.

We start with an orientable, n-dimensional (pseudo-) Riemannian manifold (M,g). The metric g induces in every point of M a nondegenerate symmetric mapping

$$g_m^{1,0} : T_mM \times T_mM \to \mathbb{R}$$
$$(X_m, Y_m) \longmapsto g_m^{1,0}(X_m, Y_m)$$

The superscript $1,0$ tells us that the function $g_m^{1,0}$ acts on tensors of type $1,0$: vectors. In general a superscript p,q attached to a function indicates that it acts on tensors of type p,q. In the sequel we will drop the subscript m for notational convenience.

The metric induces, in the same way as a symplectic form does, cf. equation (1.4.3), an isomorphism between tangent and cotangent space:

$$I^{1,0} : T_mM \to T_m^*M$$
$$\langle Y \mid I^{1,0}(X) \rangle = g^{1,0}(X,Y), \quad \forall Y \in T_mM .$$

With this isomorphism we can define the mapping

$$g^{0,1} : T_m^*M \times T_m^*M \to \mathbb{R}$$

by

$$g^{0,1}(\alpha,\beta) = g^{1,0}((I^{1,0})^{-1}(\alpha),\ (I^{1,0})^{-1}(\beta)) .$$

So the metric g gives in every point m of M not only a scalar product on the space of vectors, but also on the space of 1-forms. One can even generalize to a scalar product $g^{p,q}$ for arbitrary tensors:

$$g^{p,q} : [(T_mM)^p \otimes (T_m^*M)^q] \times [(T_mM)^p \otimes (T_m^*M)^q] \to \mathbb{R}$$

is defined as follows:

Let A and B be tensors of type p,q , then we can write with respect to some reference frame

$$A = A_{i_1 \dots i_p}^{j_1 \dots j_q} X^{i_1} \otimes \dots \otimes X^{i_p} \otimes \omega_{j_1} \otimes \dots \otimes \omega_{j_q}$$

and similarly for B . Then

$$g^{p,q}(A,B) = A_{i_1 \dots i_p}^{j_1 \dots j_p} B_{\hat{\imath}_1 \dots \hat{\imath}_p}^{\hat{\jmath}_1 \dots \hat{\jmath}_q} g^{1,0}\left(X^{i_1}, X^{\hat{\imath}_1}\right) g^{1,0}\left(X^{i_2}, X^{\hat{\imath}_2}\right) \dots$$
$$\dots g^{0,1}\left(\omega_{j_q}, \omega_{\hat{\jmath}_q}\right) .$$

Also the isomorphism can be extended to p,q tensors:

$$I^{p,q} : (T_m M)^p \otimes (T_m^* M)^q \longrightarrow (T_m^* M)^p \otimes (T_m M)^q$$

is defined by

$$\langle B \mid I^{p,q}(A) \rangle = g^{p,q}(A,B) .$$

In the language of classical tensor calculus the mapping $g^{p,q}$ contracts all corresponding indices on two tensors of type p,q and the isomorphism $I^{p,q}$ raises all p lower indices and lowers the q upper indices of a tensor of type p,q .

We remark that the mapping $I^{p,q}$ respects the symmetry or antisymmetric of the tensors they act on. For instance, $I^{o,k}$ transforms a k-form into an antisymmetric tensor of type k,o : a k-vector.

B. The Hodge star on $\Lambda^k(M)$.

We choose an orthonormal base $\{\omega_i\}$ of 1-forms on M and take the wedge product to define the *volume form* V :

$$V = \omega_1 \wedge \omega_2 \wedge \dots \wedge \omega_n , \quad \dim M = n .$$

One can prove that V is independent of the choice of the orthonormal frame, except for a sign, corresponding to a choice of orientation.

Define the *Hodge star operator* by

$$* : \Lambda^k(M) \to \Lambda^{n-k}(M)$$

$$*\alpha = \frac{1}{k!} V(I^{o,k})(\alpha)) .$$

Some comments on this definition: α is a k-form, $I^{o,k}(\alpha)$ is a k-vector and $V(I^{o,k}(\alpha))$ is the contraction of that k-vector with the volume form. The factor $\frac{1}{k!}$ is conventional.

One can show that, for $\alpha,\beta \in \Lambda^k(M)$, one has

$$\alpha\Lambda * \beta = g^{o,k}(\alpha,\beta)V = \beta\Lambda * \alpha .$$

Another property of the star operator is

$$**\alpha = (-)^{|g|}(-1)^{k(n-k)}\alpha , \quad \alpha \in \Lambda^k(M)$$

$|g|$ is the determinant of the matrix

$$g^{ij} = g^{1,0}(X^i,X^j)$$

for any basis $\{X^i\}$ of T_mM .

C. The Hodge star on $\Lambda^k(M,\Gamma(E))$.

Consider a form $\alpha_E \in \Lambda^k(M,\Gamma(E))$. Locally we can write

$$\alpha_E = \alpha \otimes f$$

with $\alpha \in \Lambda^k(M)$ and $f \in F$.

The Hodge star on $\Lambda^k(M,\Gamma(E))$ is then defined by

$$*_E : \Lambda^k(M,\Gamma(E)) \to \Lambda^{n-k}(M,\Gamma(E))$$

$$*_E\alpha_E = *\alpha \otimes f .$$

The star operator acts only on the form α and not on the vector f.

D. Hodge star on $\Lambda^k(P,F)$.

Using the isomorphism

$$Is : \Lambda^k(M,\Gamma(E)) \to \bar{\Lambda}^k(P,F)$$

discussed in section (3.2.3) we define the Hodge star operator

$$*_F : \bar{\Lambda}^k(P,F) \to \bar{\Lambda}^{n-k}(P,F)$$

by $$*_F = Is \circ *_E \circ Is^{-1} .$$

E. Metric on $\Lambda^k(M)$.

Let α and β be points of $\Lambda^k(M)$. Then for every point of M we have the n-form:

$$\alpha\wedge * \beta = g^{o,k}(\alpha,\beta)V$$

which can be integrated over M to give a metric, a bilinear symmetric non-degenerate mapping, on $\Lambda^k(M)$:

$$(\ ; \ ) : \Lambda^k(M)\times\Lambda^k(M) \to \mathbb{R}$$

$$(\alpha;\beta) = \int_M \alpha\wedge * \beta = \int_M g^{o,k}(\alpha,\beta)V = (\beta;\alpha) .$$

Here we have to assume something about α, β or M for the integral to exist. We can, for instance, take M to be compact or α and β to be of compact support etc.

F. Metric on $\Lambda^k(M,\Gamma(E))$.

To introduce a metric on $\Lambda^k(M,\Gamma(E))$ we need a metric h on F. Let $\{(U_i,\phi_i)\}$ be a trivializing atlas for $\Lambda^k(M,\Gamma(E))$. We mean by this that for a point α_E of $\Lambda^k(M,\Gamma(E))$ we have on U_i a coordinate representation of the form:

$$\phi_i(\alpha_E) = \alpha \otimes f^i_\alpha , \qquad \alpha \in \Lambda^k(M) , \qquad f^i_\alpha \in F .$$

If $U_i \cap U_j$ is not empty we have

$$\left.\begin{aligned} \phi_i(\alpha_E) &= \alpha \otimes f^i_\alpha \\ \phi_j(\alpha_E) &= \alpha \otimes f^j_\alpha \end{aligned}\right\} \quad \text{on} \quad U_i \cap U_j$$

and the relation between f^i_α and f^j_α is given by the transition function of E, a function on $U_i \cap U_j$ taking values in the representation of G on F:

$$f^i_\alpha = \mathrm{Rep}(g) f^j_\alpha$$

for some g in G (in every point of $U_i \cap U_j$).

We introduce on U_i the function

$$\overline{gh}/U_i : \Lambda^k(U_i,\Gamma(E)) \times \Lambda^k(U_i,\Gamma(E)) \to \mathbb{R}$$

by

$$\overline{gh}/U_i(\alpha_E,\beta_E) = g^{o,k}(\alpha,\beta) h(f^i_\alpha, f^i_\beta) .$$

If the metric h on F is invariant under transformations with the representations of G, the functions $\overline{gh}/U_i$ and $\overline{gh}/U_j$ will coincide on the overlap $U_i \cap U_j$ and it will be possible to define a function $\overline{gh}$ on the whole of M. This global function is used in the metric on $\Lambda^k(M,\Gamma(E))$:

$$\begin{aligned}(\alpha_E;\beta_E)_E &= \int_M \overline{gh}(\alpha_E,\beta_E)V = \int g^{o,k}(\alpha,\beta)h(f_\alpha,f_\beta)V \\ &= \int_M \alpha\Lambda * \beta)h(f_\alpha,f_\beta) .\end{aligned}$$

G. The metric on $\bar{\Lambda}^k(P,F)$.

Let α_F and β_F be points of $\bar{\Lambda}^k(P,F)$. Then the metric $(\ ;\ )_F$ on them is defined by

$$(\alpha_F;\beta_F)_F = (Is^{-1}(\alpha_F);Is^{-1}(\beta_F))_E .$$

H. Example: the metric on $\Lambda^k(M,\Gamma(\mathrm{ad}\,P))$.

In the sequel we will often use the bundle $\mathrm{ad}\,P$, the bundle associated to $P(M,G)$ and the adjoint representation of G on the Lie algebra $\ell(G)$. We assume G to be a matrix group. The adjoint action of G on the Lie algebra is then

$$\mathrm{ad} : G\times\ell(G) \to \ell(G) .$$
$$(g,\ell) \longmapsto g\ell g^{-1} .$$

We define a metric on $\ell(G)$ using the trace:

$$h : \ell(G)\times\ell(G) \to \mathbb{R}$$
$$h(\ell_1,\ell_2) = tr(\ell_1\ell_2) .$$

This metric is invariant under adjoint transformations and we define a metric on $\Lambda^k(M,\Gamma(\mathrm{ad}\,P))$ by

$$(\alpha_{\mathrm{ad}\,P};\beta_{\mathrm{ad}\,P})_{\mathrm{ad}\,P} = \int_M \alpha\Lambda * \beta\, tr(f_\alpha f_\beta)$$

with in some chart ϕ:

$$\phi(\alpha_{\mathrm{ad}\,P}) = \alpha\otimes f_\alpha$$
$$\phi(\beta_{\mathrm{ad}\,P}) = \beta\otimes f_\beta .$$

3.2.7. Codifferentials.

If one has on some linear space V a metric $(\ ;\ )$ and a linear operator A, one can introduce the adjoint operator A^+ of A by

$$(\phi;A\psi) = (A^+\phi;\psi) \qquad \forall\phi,\psi \in V .$$

As linear space we take

$$\Lambda^*(M) = \bigoplus_{p=0}^{n} \Lambda^p(M)$$

and the metric of section (3.2.6E), with forms of different degree orthogonal. We then define the adjoint δ of the exterior derivative d

$$d: \Lambda^k(M) \to \Lambda^{k+1}(M)$$

by

$$(d\alpha;\beta) = (\alpha;\delta\beta) \qquad \alpha,\beta \in \Lambda^*(M) .$$

One easily checks that δ, called the *codifferential*, lowers the degree of a form by one:

$$\delta: \Lambda^k(M) \to \Lambda^{k-1}(M) .$$

The codifferential is related to the exterior derivative and the star operator through the formula [16]

$$\delta\alpha = (-1)^{|g|}(-1)^{n(k+1)} * d * \alpha$$

with α a k-form, n the dimension of M and $|g|$ the determinant of the metric.

3.2.8. Covariant differentiation and covariant codifferentials.

To differentiate forms of type ad on $P(M,G)$ with values in $\ell(G)$ we have to introduce a generalization of the exterior derivative. This is the *covariant derivative* D_A,

$$D_A : \Lambda^k(P,\ell(G)) \to \Lambda^{k+1}(P,\ell(G))$$
$$D_A\phi = (d\phi)^H .$$

Here H is the horizontal projection on forms:

$$H: \Lambda^k(P,\ell(G)) \to \bar{\Lambda}^k(P,\ell(G))$$
$$\phi^H(X_1,\ldots,X_k) = \phi(X_1^H,X_2^H,\ldots,X_k^H)$$

with X_i^H the horizontal part (with respect to the connection A in $P(M,G)$) of the vector X_i.

One can show that for elements ϕ of $\bar{\Lambda}^k(P,\ell(G))$

$$D_A\phi = d\phi + [A \wedge \phi] .$$

Notation.

Let

$$\eta \in \Lambda^k(P,\ell(G)) , \qquad \psi \in \Lambda^m(P,\ell(G)) ,$$

then we define a product

$$[\eta \wedge \psi] \in \Lambda^{k+m}(P,\ell(G))$$

by combining the wedge product of forms and the commutator of the Lie algebra: we write locally

$$\eta = \bar{\eta} \otimes L \qquad \bar{\eta} \in \Lambda^k(M),\ L,\ K \in \ell(G)$$
$$\psi = \bar{\psi} \otimes K \qquad \bar{\psi} \in \Lambda^m(M)$$

then

$$[\eta \wedge \psi] = \bar{\eta} \wedge \bar{\psi} \otimes [L,K]$$

One can also show that the covariant derivative of the connection is

$$D_A A = dA + \tfrac{1}{2}[A \wedge A]$$

Remember that a connection is not horizontal, therefore A does not belong to $\bar{\Lambda}^k(P,\ell(G))$. The 2-form

$$F(A) = D_A A$$

is the curvature of the connection A. $F(A)$ belongs to $\bar{\Lambda}^2(P,\ell(G))$. The curvature F satisfies the Bianchi-identity $D_A F = 0$.

In section 3.2.6G,H we defined a metric $(\ ;\ )_{\ell(G)}$ on $\bar{\Lambda}^k(P,\ell(G))$ and therefore we can define the adjoint operator of D_A. This is the *covariant codifferential* δ_A. One can prove that δ_A is in the same way related to D_A and $*_{\ell(G)}$ as δ is related to d and $*$:

$$\delta_A \phi = (-1)^{g}\,(-1)^{n(k+1)} *_{\ell(G)} \circ D_A \circ *_{\ell(G)} \phi \;.$$

3.3. The Yang-Mills equation.

3.3.1. The manifestly covariant formulation.

Let M^4 be a four-dimensional Riemannian or pseudo-Riemannian manifold, $P(M^4,G)$ a principle fiber bundle over M^4 with structure group G and $C = C(P(M^4,G))$ the set of connections in P.

We call an element $A \in C$ a *Yang-Mills connection* if it is an extremum of the *Yang-Mills functional* or *Action* S:

$$S(A) = -\tfrac{1}{2}(F(A);\ F(A))_{\ell(G)} = -\tfrac{1}{2}\int_M \mathrm{tr}(F(A) \wedge *_{\ell(G)} F(A))$$

with $F(A)$ the curvature,

$$F(A) = D_A A = dA + \tfrac{1}{2}[A \wedge A],$$

$(\ ;\ )_{\ell(G)}$ the inner product of section 3.2.6 G, H and $*_{\ell(G)}$ the Hodge star of 3.2.6D. From now on we will not distinguish between Hodge stars and metrics on $\Lambda^k(M), \Lambda^k(M, \ell(E))$ and $\bar{\Lambda}^k(P,F)$ and we will accordingly drop the subscripts indicating the difference.

PROPOSITION.

A Yang-Mills connection A *satisfies*

$$\delta_A F(A) = 0 \qquad (3.3.1.1)$$

PROOF.

Since C is an affine space modelled on $\bar{\Lambda}^1(P, \ell(G))$ (see section 3.2.5), we can write for an arbitrary element of C .

$$A_t = A + t\lambda\ , \qquad \lambda \in \bar{\Lambda}^1(P, \ell(G))$$

The connection A is an extremum of the action if

$$\frac{d}{dt} S(A_t) = 0$$

for all $\lambda \in \bar{\Lambda}^1(P, \ell(G))$.

One easily calculates

$$F(A_t) = D_{A_t}(A_t) = D_A(A) + D_A\lambda + 0(t^2)$$

and

$$A(A_t) = S(A) + t(D_A\lambda; F(A)) + 0(t^2)\ .$$

Therefore

$$\frac{d}{dt} S(A_t) = (D_A\lambda ; F) = (\lambda ; \delta_A F) .$$

This is zero for all λ if and only if equation (3.3.1.1) is satisfied. ∎

Equation (3.3.1.1) is called the Yang-Mills equation. Notice that it has the same form in Euclidean and in Minkowski space-time. The metric structure is hidden in the definition of the codifferential. If one writes out the equation in components the difference will appear.

3.3.2. The 3+1 decomposition of space-time.

We want to look at the Yang-Mills equations as the equations of motion of a Hamiltonian system. In this formalism time plays a special rôle and we assume therefore that M^4 can be decomposed as

$$M^4 = \mathbb{R} \times M^3 \tag{3.3.2.1}$$

with $\mathbb{R}$ the time coordinate axis and M^3 space, a three-dimensional Riemannian space. The splitting (3.3.2.1) induces a splitting of the tangent space at every point of M^4. We write

$$X_m = a^o(t,\vec{x}) \frac{\partial}{\partial t} + a^i(t,\vec{x}) \frac{\partial}{\partial x^i}$$

with

$$(t,\vec{x}) \in M^4 , \quad X_m \in T_m M^4 , \quad \frac{\partial}{\partial x^i} \in T_{\vec{x}} M^3 .$$

We assume the metric to be of the form:

$$g\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\right) = \rho$$

$$g\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x^i}\right) = 0$$

$$g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = g_3\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) .$$

Depending on whether one studies an Euclidean theory or the Minkowski version of it, the parameter ρ is plus or minus one. g_3 is the Riemannian metric of M^3.

The splitting of space-time also extends in a natural way to forms: let ω^p be a p-form, then we write

$$\omega^p = dt \wedge \omega_N^{p-1} + \omega_T^p$$

with

$$\omega_N^{p-1} = i\left(\frac{\partial}{\partial t}\right)\omega^p , \qquad \omega_T^p = \omega^p - dt \wedge \omega_N^{p-1} .$$

This splitting is orthogonal with respect to the metric of forms discussed in section 3.2.6E: the normal part of $\omega^p : dt \wedge \omega_N^{p-1}$ contains dt while the tangent part ω_T^p does not. Therefore

$$dt \wedge \omega_N^{p-1} \wedge *\omega_T^p = 0$$

(since $*\omega_T^p$ does contain dt) and accordingly

$$(dt \wedge \omega_N^{p-1} ; \omega_T^p) = 0 .$$

The action of the exterior derivative can also be decomposed:

$$d\omega^p = d(dt \wedge \omega_N^{p-1} + \omega_T^p) = dt \wedge (\dot{\omega}_T^p - d_3\omega_N^{p-1}) + d_3\omega_T^p$$

with d_3 the exterior derivative on M^3 and

$$\dot{\omega}_T^p = \frac{d}{dt}(\omega_T^p) .$$

The inner product of forms on M^4 is decomposed as follows:

$$\begin{aligned}(\alpha;\beta) &= \int_{M^4} g^{o,k}(\alpha,\beta)\,dV_3 dt \\ &= \int_{M^4} [g^{o,k}(dt\wedge\alpha_N,\ dt\wedge\beta_N) + g^{o,k}(\alpha_T,\beta_T)]\,dV_3 dt \\ &= \int_{M^4} [g^{o,1}(dt,dt)\,g_3^{o,k-1}(\alpha_N,\beta_N) + g_3^{o,k}(\alpha_T,\beta_T)]\,dV_3 dt \\ &= \int_{\mathbb{R}} [\rho(\alpha_N;\beta_N)_3 + (\alpha_T;\beta_T)_3]\,dt \qquad (3.3.2.3)\end{aligned}$$

Here dV_3 is the volume element on M^3 and $(\ ;\ )_3$ the metric of forms on M^3, induced by g_3. We have used the orthogonality of normal and tangent components.

The results (3.3.2.3) can be generalized to form α_E, β_E with values in sections of an associated fiber bundle E, provided we can define an appropriate metric on such forms (see section 3.2.6F). First, we split α_E (and β_E similarly) in normal and tangent components by using a chart ϕ of E and writing:

$$\phi(\alpha_E) = \alpha \otimes f$$

and splitting the 1-form:

$$\alpha = dt\wedge\alpha_N + \alpha_T$$

This defines the splitting of α_E:

$$\alpha_{EN} = \phi^{-1}(dt\wedge\alpha_N \otimes f)\ , \qquad \alpha_{ET} = \phi^{-1}(\alpha_f \otimes f)$$

This splitting is independent of the choice of chart ϕ, because in another chart only f is changed, not the form α. By a calculation similar to the one leading to (3.3.2.3) we find

$$(\alpha_E;\beta_E) = \int_{\mathbb{R}} [\rho(\alpha_{EN};\beta_{EN})_{3E} + (\alpha_{ET};\alpha_{ET})_{3E}]\,dt \qquad (3.3.2.4)$$

with $(\ ;\ )_{3E}$ the metric on forms on M^3 with values in sections of the pullback bundle of R to M^3.

With the isomorphism of section 3.2.3 between $\Lambda^k(M,\ell(E))$ and $\bar{\Lambda}^k(P,F)$ we can make a 3+1 decomposition of forms in $\Lambda^k(P,F)$ and also decompose the metric of sections of $\Lambda^k(P,F)$ in the obvious way.

We use (3.3.2.4) to write for the Yang-Mills action

$$S(A) = -\tfrac{1}{2}(F(A);F(A)) = \int L(t)dt$$

We interpret the function $L(t)$ as the Lagrangian for the Yang-Mills system. To calculate $L(t)$ we have to decompose $F(A)$. Because we want to express everything in terms of the connection A, we have also to decompose A. Here we encounter a slight difficulty. F is a horizontal form (of type ad) on P and as such isomorphic to a form on M. We can then split this form on M and return to P by the isomorphism: this defines uniquely the splitting:

$$F = F_N + F_T$$

However, A is not horizontal and therefore not naturally isomorphic to a form on M. To get a form on M we choose a section

$$s: M \to P$$

and pull A back to M by this section

$$A_M = s^*A.$$

We can split A_M in normal and tangent parts

$$A_M = dt \wedge A_N + A_T. \qquad (3.3.2.5)$$

We then write for $F(A)$:

$$F(A) = dt \wedge (\dot{A}_T - D^3_{A_T}A_N) + D^3_{A_T}A_T \qquad (3.3.2.6)$$

with $D^3_{A_T}$ the covariant derivative on $P(M^3,G)$ and

$$D^3_{A_T} A_N = d_3 A_N + [A_T \wedge A_N]$$
$$D^3_{A_T} A_T = d_3 A_T + \tfrac{1}{2}[A_T \wedge A_T]$$

We leave it to the reader to check that though the splitting of A (3.3.2.5) is not independent of the choice of section s, the splitting of F (3.3.2.6) is independent.

If we now use (3.3.2.4) we find for L:

$$L(t) = -\tfrac{1}{2}\rho(A_T - D^3_{A_T} A_N ; \dot{A}_T - D^3_{A_T} A_N)_3 - \tfrac{1}{2}(D_{A_T} A_T ; D_{A_T} A_T)_3 .$$

In the rest of this paper we will use only the 3+1 decomposed formalism, so it is convenient to change the notation by dropping sub and superscripts 3 and introduce

$$A = A_T , \quad A_o = A_N , \quad D_A = D^3_{A_T} .$$

With these conventions the Lagrangian becomes

$$L(t) = -\tfrac{1}{2}\rho(\dot{A}-D_A A_o ; \dot{A}-D_A A_o) - \tfrac{1}{2}(D_A A ; D_A A) . \qquad (3.3.2.7)$$

3.3.3. The Yang-Mills system as a constrained Hamiltonian system.

Motivated by the form (3.3.2.7) of the Lagrangian that we derived from the action S(A) we take as a configuration space for the Yang-Mills system a space Q consisting of pairs (A_o, A) with

$$A_o \in \bar{\Lambda}^0(P_3, \ell(G)) , \quad A \in C(P_3) .$$

$P_3 = P_3(M^3,G)$ is the principle fiber bundle obtained from $P(M^4,G)$ by pullback with the submanifold map j :

$$j : M^3 \to M^4 = \mathbb{R} \times M^3$$

and

$$P_3 = j^*P(M^4,G) .$$

Notice that the mapping j is time dependent, so P_3, A_o and A are also dependent on time.

We have seen in section (3.2.5) that the tangent space to $C(P_3)$ is $\bar{\Lambda}^1(P_3,\ell(G))$. Therefore velocity phase space is:

$$TQ = Q \times \bar{\Lambda}^0(P_3,\ell(G)) \times \bar{\Lambda}^1(P_3,\ell(G)) \qquad (3.3.3.1)$$

We assume that the metrics on $\bar{\Lambda}^0(P_3,\ell(G))$ and $\bar{\Lambda}^1(P_3,\ell(G))$ are nondegenerate. Since M^3 is a Riemannian manifold these metrics are nondegenerate if and only if the Lie algebra metric is nondegenerate. By the Cartan criterium this is equivalent to demanding that the Lie algebra is semi-simple, which we do from now on. If these metrics (on $\bar{\Lambda}^k(P_3,\ell(G))$) are nondegenerate we can identify according to the Riesz lemma, the space $\bar{\Lambda}^k(P_3,\ell(G))$ and its dual. We therefore use the metric for the contraction between T^*Q and TQ. More explicitly, let (A_o,A) be a point of Q, $X = (A_o,A,\dot{A}_o,\dot{A})$ a point of TQ and $\omega = (A_o,A,\pi_o,\pi)$ a point of T^*Q. Then the contraction of ω and X is:

$$\langle\omega \mid X\rangle = \int tr(\pi_o \wedge * \dot{A}_o) + \int tr(\pi \wedge * A) = (\pi_o;\dot{A}_o) + (\pi;A) .$$

Notation.

It is useful to introduce the following notation: Let X and ω be as above, then we write them as:

$$X = \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A} \frac{\delta}{\delta A} , \quad \omega = \pi_o \delta A_o + \pi \delta A$$

The symbols $\frac{\delta}{\delta A_o}$, $\frac{\delta}{\delta A}$ etc. are a reminder of to which space the different components of vectors and forms belong. With this notation calculations become quite similar to calculations in finite dimensions. For instance,

$$\langle \omega \mid X \rangle = \langle \pi_o \delta A_o + \pi \delta A \mid \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A} \frac{\delta}{\delta A} \rangle$$

$$= \langle \pi_o \delta A_o \mid \dot{A}_o \frac{\delta}{\delta A_o} \rangle + \langle \pi \delta A \mid \dot{A} \frac{\delta}{\delta A} \rangle$$

$$= (\pi_o ; \dot{A}_o) + (\pi ; \dot{A}) .$$

This looks like in the finite dimensional case:

$$\bar{X} = a^i \frac{\partial}{\partial x^i} + b^i \frac{\partial}{\partial y^i} , \quad \bar{\omega} = c_i dx^i + d_i dy^i$$

$$\langle \bar{\omega} \mid \bar{X} \rangle = \langle a^i \frac{\partial}{\partial x^i} + b^i \frac{\partial}{\partial y^i} \mid c_i dx^i + d_i dy^i \rangle$$

$$= a^i c_i + b^i d_i .$$

The summation over the index i in the finite dimensional case is replaced by an integration. A vector $X = \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A} \frac{\delta}{\delta A}$ acts on a function $f(A_o, A)$ according to:

$$X(f) = \frac{d}{d\lambda} f(A_o + \lambda \dot{A}_o , A + \lambda \dot{A}) = \frac{d}{d\lambda} f(A_o + \lambda \dot{A}_o , A) + \frac{d}{d\lambda} f(A_o , A + \lambda \dot{A}) \tag{3.3.3.2}$$

A commutator of vector fields is calculated in the familiar way

$$X = f_o(A_o, A) \frac{\delta}{\delta A_o} + f(A_o, A) \frac{\delta}{\delta A} , \quad Y = g_o(A_o, A) \frac{\delta}{\delta A_o} + g(A_o, A) \frac{\delta}{\delta A}$$

$$[X,Y] = X(g_o) \frac{\delta}{\delta A_o} + X(g) \frac{\delta}{\delta A} - Y(f_o) \frac{\delta}{\delta A_o} - Y(f) \frac{\delta}{\delta A} \tag{3.3.3.3}$$

with $X(g_o)$ etc. calculated according to equation (3.3.3.2)

We now want to go from velocity phase space TQ to phase space T^*Q, using the fiber derivative. Let X and Y be two vector fields on Q. We write

$$X = \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A} \frac{\delta}{\delta A} , \quad Y = \dot{B}_o \frac{\delta}{\delta A_o} + \dot{B} \frac{\delta}{\delta A} , \quad FL(X) = \pi_o \delta A_o + \pi \delta A .$$

With equation (1.2.2):

$$<Y \mid FL(X)> = <\dot{B}_o \frac{\delta}{\delta A_o} + \dot{B}\frac{\delta}{\delta A} \mid \pi_o \delta A_o + \pi \delta A> =$$

$$= \frac{d}{d\lambda} L(A_o, A, \dot{A}_o + \lambda \dot{B}_o, \dot{A} + \lambda \dot{B}) .$$

Using the form (3.3.2.7) of the Lagrangian we obtain:

$$(\dot{B}_o, \pi_o) + (\dot{B}; \pi) = -\rho(\dot{B}; \dot{A} - D_A A_o) .$$

As $\dot{B}_o$ and $\dot{B}$ are arbitrary we find for the components of $FL(X)$:

$$\pi_o = 0$$

$$\pi = -\rho(\dot{A} - D_A A_o) .$$

The range of FL is a submanifold M_1 of T^*Q on which

$$\pi_o = 0 .$$

On M_1 we define the Hamiltonian as in equation (1.2.3):

$$H \circ FL(X) = <X \mid FL(X)> - L(X)$$

$$= - (\dot{A}; \dot{A} - D_A A_o) + \tfrac{1}{2} (\dot{A} - D_A A_o; A - D_A \dot{A}_o) + \tfrac{1}{2}(D_A A; D_A A)$$

$$= -\rho(\pi;\pi) + (\pi; D_A A_o) + \tfrac{1}{2}(D_A A; D_A A) \qquad (3.3.3.4)$$

To define Hamiltonian mechanics on M_1 we need a symplectic form on M_1 . The natural candidate is

$$\omega_1 = j_1^* \omega ,$$

the pullback of the symplectic form T^*Q by the submanifold map

$$j_1 : M_1 \to T^*Q .$$

However, we have the following

PROPOSITION 3.3.3.1.

ω_1 *is not symplectic* .

PROOF.

We will show ω_1 is degenerate. In Chapter I (equation (1.6.9)) we calculated ω :

$$\omega(Y_1,Y_2) = \langle e_2 \mid \beta_1 \rangle - \langle e_1 \mid \beta_2 \rangle$$

with

$$Y_1 = (x,\alpha,e_1,\beta_1) \qquad Y_2 = (x,\alpha,e_2,\beta_2) .$$

In the present notation we have

$$x = (A_o,A)$$

$$\alpha = (\pi_o,\pi)$$

$$e_1 = \dot{A}_{o1}\frac{\delta}{\delta A_o} + \dot{A}_1\frac{\delta}{\delta A} , \qquad e_2 = \dot{A}_{o2}\frac{\delta}{\delta A_o} + \dot{A}_2\frac{\delta}{\delta A}$$

$$\beta_1 = \dot{\pi}_{o1}\frac{\delta}{\delta\pi_o} + \dot{\pi}_1\frac{\delta}{\delta\pi} , \qquad \beta_2 = \dot{\pi}_{o2}\frac{\pi}{\delta\pi_o} + \dot{\pi}_2\frac{\delta}{\delta\pi} .$$

So now equation (1.6.9) reads

$$(Y_1,Y_2) = (\dot{\pi}_{o1};\dot{A}_{o2}) + (\dot{\pi}_1;\dot{A}_2) - (\dot{\pi}_{o2};\dot{A}_{o1}) - (\dot{\pi}_2;\dot{A}_1) .$$

Let X and Y be vectors in TM_1 , then

$$\omega_1(\widetilde{X},\widetilde{Y}) = \omega(j_*\widetilde{X}, j_*\widetilde{Y}) .$$

The vectors $j_*\widetilde{X}$ and $j_*\widetilde{Y}$ do not contain components in the $\frac{\delta}{\delta\pi_o}$ direction. Identifying $j_*\widetilde{X}$ and $\widetilde{X}$, and $j_*\widetilde{Y}$ and $\widetilde{Y}$ we find for ω_1

$$\omega_1(\widetilde{X},\widetilde{Y}) = (\dot{\pi}_{\widetilde{X}};\dot{A}_{\widetilde{Y}}) - (\dot{\pi}_{\widetilde{Y}};\dot{A}_{\widetilde{X}}) \qquad (3.3.3.5)$$

It will now be clear that ω_1 is degenerate, because any vector in the $\frac{\delta}{\delta A_o}$ direction gives zero contribution to ω_1. So ω_1 is not symplectic. ∎

The conclusion is that (M_1,ω_1,H) is a presymplectic system and we have to apply the constraint algorithm.

3.4. The constraint algorithm for the Yang-Mills system.

We want to find a solution of

$$i(X_H)\omega_1 = -dH \tag{3.4.1}$$

on M_1, with

$$H = -\tfrac{1}{2}\rho(\pi;\pi) + (\pi;D_A A_o) + \tfrac{1}{2}(D_A A;D_A A) . \tag{3.4.2}$$

In chapter two we found that the condition for a solution to exist on M_1 was:

$$\langle TM_1^{\perp} \mid dH \rangle = 0 . \tag{3.4.3}$$

PROPOSITION 3.4.1.

The set of points of M_1 *for which* (3.4.3) *is true is the submanifold* M_2 *of* M_1 *defined by*

$$M_2 = \{(A_o,A,\pi) \in M_1 \mid \delta_A \pi = 0\} .$$

PROOF.

We have

$$TM_1^{\perp} = \{X \in TM_1 \mid \omega_1(X,Y) = 0 , \ \forall Y \in TM_1\}$$

$$= \{\dot{A}_o \frac{\delta}{\delta A_o} \mid \dot{A}_o \in \bar{\Lambda}^o(P_3,\ell(G))\}$$

Using this

$$<\dot{A}\frac{\delta}{\delta A_o} \mid dH> = \frac{d}{dt}H(A_o + t\dot{A}_o)$$

$$= +(\pi;D_A\dot{A}_o)$$

Since $\dot{A}_o$ is arbitrary this expression is zero iff

$$\delta_A\pi = 0 . \blacksquare$$

On M_2 one has solutions of (3.4.1). These solutions are tangent to M_2 iff

$$<M_2^{\perp} \mid dH> = 0 . \tag{3.4.4}$$

PROPOSITION 3.4.2.

For all points of M_2 equation (3.4.4) is satisfied.

PROOF.

We have to find the form of $TM_2^{\perp}$. To this end we first prove some lemmata.

LEMMA 1.

Let $Y = \dot{A}_o\frac{\delta}{\delta A_o} + \dot{A}\frac{\delta}{\delta A} + \dot{\pi}\frac{\delta}{\delta\pi}$

be a vector in TM_1/M_2. Then

$$Y \in \underline{TM}_2 \Longleftrightarrow (\dot{\pi};D_A\eta) - (\dot{A};[\pi\ \eta]) = 0 .$$

PROOF of lemma 1.

Since all functions of the form

$$f_\eta(A,\pi) = (\pi;D_A\eta) , \quad \eta \in \bar{\Lambda}^o(P_3,\ell(G))$$

are zero on M_2, we have

$$Y \in \underline{TM}_2 \Longleftrightarrow Y(f_\eta) = 0 \qquad \forall f_\eta .$$

But

$$Y(f_\eta) = \frac{d}{d\lambda}(\pi + \lambda\dot{\pi}; D_{A+\lambda\dot{A}}\eta) = (\dot{\pi}; D_A\eta) + (\pi; [\dot{A},\eta]) .$$

We rewrite the last term as

$$(\pi; [\dot{A},\eta]) = -(\dot{A}; [\pi,\eta]) \qquad (3.4.5)$$

This formula is based on the fact that the trace used to define the metric of Lie algebra valued forms is cyclic: for arbitrary $\alpha,\beta,\gamma \in \ell(G)$ we have

$$tr(\alpha[\beta,\gamma]) = tr(\alpha\beta\gamma - \alpha\gamma\beta) = tr(\alpha\beta\gamma - \beta\alpha\gamma) = tr([\alpha,\beta]\gamma) .$$

This proofs Lemma 1. ▼

<u>LEMMA 2.</u>

If $$Y = \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A}\frac{\delta}{\delta A} + \dot{\pi}\frac{\delta}{\delta\pi} \in \underline{TM}_2$$

and

$$(\dot{\pi},B) - (\dot{A},C) = 0$$

then

$$B = D_A\eta \quad , \quad C = [\pi,\eta]$$

for some

$$\eta \in (P_3,\ell(G)).$$

PROOF of lemma 2.

Define a function

$$\widetilde{\pi A} : \bar{\Lambda}^{o}(P_3,\ell(G)) \to \bar{\Lambda}^{1}(P_3,\ell(G)) \times \bar{\Lambda}^{1}(P_3\ell(G))$$

by

$$\eta \longmapsto (D_A\eta,\ -[\pi,\eta])\ .$$

In Proposition 2.2.5 we proved

$$(S^{\vdash})^{\vdash} = \bar{S}$$

for any subspace S of a Banach space E. Take E to be the product $\bar{\Lambda}^{1}(P_3,\ell(G)) \times \bar{\Lambda}^{1}(P_3,\ell(G))$ and S the image of $\widetilde{\pi A} : S = \widetilde{\pi A}(\bar{\Lambda}^{o}(P_3,\ell(G))$.
By Lemma 1 we have

$$S^{\vdash} = \{(\dot{\pi},\dot{A}) \in \bar{\Lambda}^{1}(P_3,\ell(G)) \times \bar{\Lambda}^{1}(P_3,\ell(G)) \mid (\dot{\pi};D_A\eta) - (\dot{A};[\pi,\eta]) = 0\}$$

$$= \{(\dot{\pi},\dot{A}) \mid \dot{A}_o \frac{\delta}{\delta A_o} + \dot{A}\frac{\delta}{\delta A} + \dot{\pi}\frac{\delta}{\delta\pi} \in \underline{TM}_2\}\ .$$

We identify here $\bar{\Lambda}^{k}(P_3,\ell(G))$ and its dual.

Now let $Y \in \underline{TM}_2$, then $(\dot{\pi},\dot{A}) \in S^{\vdash}$. Furthermore, let $(\dot{\pi};B) - (\dot{A},C) = 0$.
This means that $(B,-C) \in (S^{\vdash})^{\vdash}$ or by Proposition 2.2.5, $(B,-C) \in S$, since S is closed. We find

$$B = D_A\eta\ ,\qquad C = [\pi,\eta] \quad \text{for some} \quad \eta \in \bar{\Lambda}^{o}(P_3,\ell(G)).\qquad \blacktriangledown$$

COROLLARY.

$$TM_2^{\perp} = A_o\frac{\delta}{\delta A_o} + [\pi,\eta]\frac{\delta}{\delta\pi} + D_A\eta\frac{\delta}{\delta A}\ .$$

PROOF of the corollary.

$$X \in TM_2^\perp \Longleftrightarrow \omega_1(X,Y) = 0 \qquad \forall Y \in \underline{TM}_2$$

$$\Longleftrightarrow (\dot{\pi}_X; \dot{A}_Y) - (\dot{\pi}_Y; \dot{A}_X) = 0 .$$

Applying Lemma 2 leads to

$$\dot{\pi}_X = [\pi\ \eta] , \qquad \dot{A}_X = D_A\eta . \quad \blacktriangledown$$

We now start the proof of the proposition by calculating, using the form (3.4.2) of the Hamiltonian:

$$< TM_2^\perp \mid dH > = \frac{d}{d\lambda} H(A_o+\lambda\dot{A}_o,\ A+\lambda\dot{A},\ \pi+\lambda\dot{\pi})$$

$$= -(\pi; D_A\dot{A}_o) + (\pi; [D_A\eta \wedge A_o]) + (D_A D_A\eta; D_A A)$$

$$-\rho([\pi,\eta];\pi) + ([\pi,\eta]; D_A A_o) .$$

The first term is zero:

$$(\pi; D_A\dot{A}_o) = 0 ,$$

because we are on M_2. Collecting terms with A_o we find

$$(\pi;[D_A\eta \wedge A_o]) + ([\pi,\eta]; D_A A_o) = (\pi; [D_A\eta \wedge A_o]) + (\pi; [\eta \wedge D_A A_o]) =$$

$$= (\pi; D_A[\eta \wedge A_o]) = 0 .$$

Here we used property (3.4.5) of the inner product, the fact that $[\eta \wedge A_o]$ is an element of $\bar{\Lambda}^o(P_3, \ell(G))$ and again that we are on M_2.

To show that the other terms:

$$\rho([\pi,\eta];\pi) , \qquad (D_A D_A\eta; D_A A)$$

are zero we need the following property of the inner product:

$$(\alpha;\beta) = (e^{\lambda\eta}\alpha\, e^{-\lambda\eta};\ e^{\lambda\eta}\beta\, e^{-\lambda\eta}) , \quad \forall\eta \in \bar{\Lambda}^{o}(P_3,\ell(G)) . \tag{3.4.6}$$

The proof of this is based on the fact that the trace is cyclic:

$$\mathrm{tr}(e^{\lambda\eta}\alpha\, e^{-\lambda\eta}\, e^{\lambda\eta}\beta\, e^{-\lambda\eta}) = \mathrm{tr}(\alpha\beta) .$$

Applying this to $\alpha = \beta = \pi$ and differentiating with respect to λ we get

$$0 = \frac{d}{d\lambda}(e^{\lambda\eta}\pi\, e^{-\lambda\eta};\ e^{\lambda\eta}\pi\, e^{-\lambda\eta}) = 2([\eta,\pi];\pi) = -2([\pi,\eta];\pi).$$

The only remaining term (of equation (3.4.4)) is

$$(D_A D_A \eta,\ D_A A).$$

We can induce a transformation

$$D_A A \longmapsto e^{\lambda\eta} D_A A\ e^{-\lambda\eta}$$

by making a transformation

$$A \longmapsto \hat{A}_\lambda = e^{\lambda\eta} A\, e^{-\lambda\eta} + e^{\lambda\eta} d\, e^{-\lambda\eta} .$$

Using this in (3.4.6) and differentiating we find

$$(D_A D_A \eta;\ D_A A) = 0 .$$

This proves the proposition. ∎

This rather lengthly calculation shows that M_2 is the final constraint manifold.

3.5. Gauge freedom for the Yang-Mills equations. The equations of motion.

We summarize the results of the first sections of this chapter.

We have decomposed space-time M^4 into the product of a time-axis and a space manifold:

$$M^4 = \mathbb{R} \times M^3 .$$

On M^4 we have a principle fiber bubdle $P_4(M^4,G)$ which induces, after splitting space-time, a bundle $P_3(M^3,G)$ on M^3. The dynamical objects of Yang-Mills theory are connections on P_4. In the 3+1 decomposed form these become elements of the space Q of pairs (A_o,A) with A_o a (time dependent) Lie algebra valued function of type ad on P_3 and A a (time dependent) connection form in P_3. The cotangent bundle of Q consists of quadrupples (A_o,A,π_o,π). The dynamics takes place in the $\pi_o = 0$ constraint manifold M_1 of T^*Q. M_1 consists of triples (A_o,A,π). On M_1 we have the presymplectic form $\omega_1 = j_1^*\omega$, with $j_1: M_1 \to T^*Q$ the embedding of M_1 and ω the canonical symplectic form on T^*Q. The Hamiltonian on M_1 reads:

$$H(A_o,A,\pi) = -\tfrac{1}{2}\rho(\pi;\pi) + (\pi; D_A A_o) + \tfrac{1}{2}(D_A A;D_A A) \qquad (3.5.1)$$

with $\rho = 1$ if M^4 is Euclidean and $\rho = -1$ if M^4 is Minkowski. The metric of forms $(\; ; \;)$ is defined by using the trace in the Lie algebra and the Hodge star on M_3.

To the presymplectic system (M_1,ω_1,H) we applied the constraint algorithm and we found that the manifold $M_2 \subset M_1 \subset T^*Q$ defined by

$$\pi_o = 0 , \qquad \delta_A\pi = 0$$

is the final constraint manifold. This means that the equation of motion:

$$i(X_H)\omega_1\Big|_{M_2} = -dH\Big|_{M_2} \qquad (3.5.2)$$

have at least one solution tangent to M_2. This was the position we reached at the end of section 3.4.

We now look in more detail at the final constraint manifold M_2.

PROPOSITION 3.5.1.

M_2 *is a first class submanifold.*

PROOF.

We have to show that

$$TM_2^{\perp} \subset TM_2 .$$

The general form of $TM_2^{\perp}$ is (cf. the corollary of section 3.4)

$$TM_2^{\perp} = \{\dot{A}_o \frac{\delta}{\delta A_o} + [\pi,\eta]\frac{\delta}{\delta\pi} + D_A\eta\frac{\delta}{\delta A} \mid \dot{A}_o, \eta \in \bar{\Lambda}^o(P_3,\ell(G))\} .$$

In the same section (cf. Lemma 1, section 3.4) we proved that

$$X \in \underline{TM}_2 \iff (\dot{\pi}_X, D_A\pi) - (\dot{A}_X, [\pi,\eta]) = 0 , \qquad \forall\eta .$$

So we have to show that if $X \in TM_2^{\perp}$, that is if

$$\dot{\pi}_X = [\pi,\eta] , \qquad \dot{A}_X = D_A\eta$$

this condition is satisfied. We calculate:

$$([\pi,\eta'], D_A\eta) - (D_A\eta', [\pi,\eta]) =$$

$$= (\pi, [\eta' \wedge D_A\eta) + (\pi, [D_A\eta' \wedge \eta]) = (\pi, D_A[\eta' \wedge \eta]) .$$

This is indeed zero, because $[\eta' \wedge \eta]$ belongs to $\bar{\Lambda}^o(P_3,\ell(G))$ and we are on M_2 . ∎

Next we investigate the gauge freedom on M_2. As we argued in section 2.6 we can add to a solution X_H of the equations of motion (3.5.2) any vector $Z \in G_\ell$, with

$$G_\ell = G_{\ell-1} + [G_{\ell-1},X_H] + [G_{\ell-1},G_{\ell-1}] \qquad \ell = 2,3,4, \ldots$$

with

$$G_1 = \ker\Omega_1 \cap \underline{TM}_2$$

to obtain a gauge equivalent vector field $X_H + Z$. In Appendix A it is shown that

$$G_\ell \subset TM_2^\perp .$$

In other words $TM_2^\perp$ is an upper bound for G_ℓ. The following proposition shows that the bound is saturated for $\ell = 2$ and consequently for all ℓ larger than 2.

PROPOSITION 3.5.2.

$$G_2 = TM_2^\perp .$$

PROOF.

Using

$$G_1 = TM_1^\perp = \{\dot{A}_o \frac{\delta}{\delta A_o} \mid \dot{A}_o \in \bar{\Lambda}^o(P_3, \ell(G))\} \tag{3.5.3}$$

we have

$$G_2 = TM_1^\perp + [TM_1^\perp, X_H] + [TM_1^\perp, TM_1^\perp] .$$

Because of (3.5.3) $TM_1^\perp$ is the tangent bundle of the A_o-manifold and therefore

$$[TM_1^\perp, TM_1^\perp] = TM_1^\perp .$$

We find

$$G_2 = TM_1^\perp + [TM_1^\perp, X_H] .$$

LEMMA 3.5.3.

$$X_H = (-\rho\pi + D_A A_o) \frac{\delta}{\delta A} + ([\pi, A_o] - \delta_A D_A A) \frac{\delta}{\delta \pi}$$

is a solution of the equation of motion (3.5.2).

PROOF.

We contract equation (3.5.2) with an arbitrary vector Y :

$$\omega_1(X_H, Y) = -dH(Y) . \tag{3.5.4}$$

In section 3.3.3 (see the Proof of proposition 3.3.3.1) we showed that

$$\omega_1(X_H, Y) = (\dot{\pi}_{X_H}; \dot{A}_Y) - (\dot{\pi}_Y; \dot{A}_{X_H})$$

with

$$X_H = \dot{A}_{o\,X_H} \frac{\delta}{\delta A_o} + \dot{A}_{X_H} \frac{\delta}{\delta A} + \dot{\pi}_{X_H} \frac{\delta}{\delta \pi}$$

$$Y = \dot{A}_{o\,Y} \frac{\delta}{\delta A_o} + \dot{A}_Y \frac{\delta}{\delta A} + \dot{\pi}_Y \frac{\delta}{\delta \pi} .$$

Using the expression (3.5.1) for the Hamiltonian we calculate the right-hand side of (3.5.4):

$$dH(Y) = Y(H) = \frac{d}{d\lambda} H(A_o + \lambda \dot{A}_{o\,Y},\ A + \lambda \dot{A}_Y,\ \pi + \lambda \dot{\pi}_Y) .$$

We consider the various terms separately:

$$\frac{d}{d\lambda} H(A_o + \lambda A_{o\,Y}) = (\pi; D_A A_{o\,Y}) ,$$

$$\frac{d}{d\lambda} H(A + \lambda \dot{A}_Y) = (\pi; [\dot{A}_Y, A_o]) + (D_A \dot{A}_Y; D_A A)$$

$$= (\dot{A}_Y; -[\pi, A_o] + \delta_A D_A A) ,$$

$$\frac{d}{d\lambda} H(\pi + \lambda \dot{\pi}_Y) = (\dot{\pi}_Y; -\rho\pi + D_A A_o) .$$

So

$$-dH(Y) = -(\dot{A}_{o\,Y}; \delta_A \pi) + (\dot{A}_Y; [\pi, A_o] - \delta_A D_A A)$$

$$-(\dot{\pi}_Y; -\rho\pi + D_A A_o) .$$

Substituting this in (3.5.4) and comparing terms depending on the arbitrary $\dot{A}_{oY}$, $\dot{A}_Y$, $\dot{\pi}_Y$, we find sufficient conditions for a solution of 3.5.4 to be:

1) $\delta_A \pi = 0$

2) $\dot{\pi}_{X_H} = [\pi, A_o] - \delta$

3) $\dot{A}_{X_H} = -\rho\pi + D_A A_o$

Condition 1) is by definition true on M_2, 2) and 3) give components of X_H, the $\dot{A}_{oX}$ being arbitrary. This proves the lemma. ▼

We stress that the Hamiltonian vector field of Lemma 3.5.3 is just <u>a</u> solution of the equations of motion and certainly not the unique solution.

<u>LEMMA 3.5.4.</u>

$$[TM_1^{\perp}, X_h] = D_A \eta \frac{\delta}{\delta A} + [\pi,\eta] \frac{\delta}{\delta\pi}, \; \eta \in \bar{\Lambda}^o(P_3, \ell(G)) \; .$$

<u>PROOF of lemma 3.5.4.</u>

An arbitrary element of $TM_1^{\perp}$ can be written an

$$\eta \frac{\delta}{\delta A_o} .$$

We use X_H of the form of Lemma (3.5.3). With the expression (3.3.3.3) for the commutator of vector fields, we obtain

$$[TM_1^{\perp}, X_H] = \eta \frac{\delta}{\delta A_o}(-\rho\pi + D_A A_o) \frac{\delta}{\delta A} + \eta \frac{\delta}{\delta A_o}([\pi, A_o] - \delta_A D_A A) \frac{\delta}{\delta\pi}$$

with

$$\eta \frac{\delta}{\delta A_o}(D_A A_o) = D_A \eta$$

$$\eta \frac{\delta}{\delta A_o}[\pi, A_o] = [\pi,\eta]$$

we find

$$[TM_1^{\perp}, X_H] = D_A\eta \frac{\delta}{\delta A} + [\pi,\eta]\frac{\delta}{\delta\pi} \ . \quad \blacktriangledown$$

Combining Lemma 3.5.4 and equation (3.5.3) gives:

$$G_2 = \dot{A}_o \frac{\delta}{\delta A_o} + D_A\eta \frac{\delta}{\delta A} + [\pi,\eta]\frac{\delta}{\delta\pi}$$

with $\dot{A}_o$ and η arbitrary. But this is exactly the form we found in section 3.4 for an element of $TM_2^{\perp}$: therefore

$$G_2 = TM_2^{\perp} . \quad \blacksquare$$

The upshot of this proposition is that <u>all</u> elements of $TM_2^{\perp}$ are gauge vectors. We can therefore give the general solution of the equations of motion by adding to a special solution, for instance, the one we found in Lemma 3.5.3, an arbitrary element of $TM_2^{\perp}$. In this way we obtain

$$X_H = \dot{A}_o \frac{\delta}{\delta A_o} + [-\rho\pi + D_A(A_o+\eta)]\frac{\delta}{\delta A} + \left[[\pi, A_o+\eta] - \delta_A D_A A\right]\frac{\delta}{\delta\pi} \ .$$

We can simplify by noting that if η is arbitrary also $A_o + \eta$ is, and we can incorporate A_o in η to get

$$X_H = \dot{A}_o \frac{\delta}{\delta A_o} + [-\rho\pi + D_A\eta]\frac{\delta}{\delta A} + \left[[\pi,\eta] - \delta_A D_A A\right]\frac{\delta}{\delta\pi} \ . \qquad (3.5.5)$$

3.6. Reduction of the final constraint manifold.

In the last section we found that the most general Hamiltonian vector field on M_2 (equation (3.5.5)) depends on two arbitrary functions $\dot{A}_o$ and η. To eliminate this gauge freedom we apply the reduction procedure sketched in section 2.6.

We first note that two points of M_2 of the form (A_o, A, π) and

(A_o', A, π) (i.e. differing only in their A_o component) are gauge equivalent. They can be connected by an integral curve of a gauge vector field $\dot{A}_o \frac{\delta}{\delta A_o}$, for take $\dot{A}_o = A_o' - A_o$. Then we define a curve

$$A(\lambda) = A_o + \lambda(A_o' - A_o)$$

with $A(0) = A_o$, $A(1) = A_o'$ and $\frac{d}{d\lambda} A(\lambda) = \dot{A}_o$.

This means that the A_o part of M_2 is physically irrelevant and we can without loss of generality take A_o to be zero. We then get a new manifold $\bar{M}_2$ of pairs (A, π) still satisfying the constraint

$$\delta_A \pi = 0 .$$

$\bar{M}_2$ is a submanifold of the cotangent bundle T^*C, with C the space of connections of the bundle $P_3(M^3, G)$.

On $\bar{M}_2$ we have a Hamiltonian $\bar{H}$ (compare with equation (3.3.3.4)):

$$\bar{H}(A, \pi) = -\tfrac{1}{2}\rho(\pi, \pi) + \tfrac{1}{2}(D_A A, D_A A) \tag{3.6.1}$$

and the equations of motion

$$i(X_{\bar{H}})\omega \Big|_{\bar{M}_2} = -d\bar{H}\Big|_{\bar{M}_2} \tag{3.6.2}$$

have the general solution (compare with (3.5.10))

$$X_{\bar{H}} = (-\rho\pi + D_A \eta) \frac{\delta}{\delta A} + ([\pi, \eta] - \delta_A D_A A) \frac{\delta}{\delta \pi} \tag{3.6.3}$$

with arbitrary η. Gauge vectors on $\bar{M}_2$ are:

$$X_\eta = D_A \eta \frac{\delta}{\delta A} + [\pi, \eta] \frac{\delta}{\delta \pi} . \tag{3.6.4}$$

By going from M_2 to $\bar{M}_2$ we have already eliminated some gauge freedom. Now there is only one arbitrary parameter left, viz. η.

We will show that the remaining gauge freedom on $\bar{M}_2$ is intimately related to the geometry of $P_3(M^3,G)$. To explain this we have to make a digression on "gauge" freedom in principle fiber bundles. A priori this concept of "gauge" freedom (in the context of fiber bundles) has nothing to do with the concept of gauge freedom introduced in relation with constrained Hamiltonian systems. However, the two concepts turn out to be the same for the Yang-Mills system.

We define an (active) "gauge" transformation of a principle fiber bundle P to be a smooth mapping

$$f: P \to P$$

such that

1) f is *equivariant:*

$$f(pg) = f(p)g \qquad \forall p \in P, \quad \forall g \in G$$

2) f *induces the identity* on the base space:

$$\pi(f(p)) = \pi(p)$$

(For the relationship between this definition of "gauge" transformation and the more usual (passive) one, where "gauge" freedom is related to the freedom of choice of a section to coordinatize the bundle, see Choquet-Bruhat et al. [32] p.405.)

Because of 1) and 2) every f can be written uniquely as

$$f(p) = p\tau(p)$$

with $\tau \in \bar{\Lambda}^{o}(P,G)$, the set of G valued functions on P satisfying:

$$\tau(pg) = g^{-1}\tau(p)g\,.$$

Therefore there is an isomorphism between the set of "gauge" transformations, denoted by GA(P) and $\bar{\Lambda}^o(P,G)$. The set GA(P) is a group. The multiplication is defined by composition; if $f_1(p) = p\tau_1(p)$ and $f_2(p) = p\tau_2(p)$ then $f_2 \circ f_1(p) = p\tau_1(p)\tau_2(p)$. The new function $f_2 \circ f_1$ satisfies 1) and 2). The inverse of $f_1(p)$ is defined by

$$f_1^{-1}(p) = p[\tau_1(p)]^{-1} .$$

It also satisfies 1) and 2). In fact one can prove (modulo some technical details [22]) that GA(P) is an (infinite dimensional) Lie group, with as its Lie algebra $\bar{\Lambda}^o(P,\ell(G))$. This means that we can write every $\tau \in \bar{\Lambda}^o(P,\ell(G))$ in exponential form:

$$\tau(p) = \exp[\eta(p)]$$

for some $\eta \in \bar{\Lambda}^o(P, (G))$. (See also Bleeker [16]).

A "gauge" transformation f induces a transformation $\bar{f}$ on the space of connections:

$$\bar{f}: C \to C , \qquad \bar{f}(A) = f^*(A) .$$

In this way we obtain a group $\overline{GA(P)}$ of transformations of C .

Consider a 1-parameter family of gauge transforms:

$$f_\eta^\lambda(p) = p \exp[\lambda\eta (p)] .$$

For $\lambda = 0$ this transformation is the identity. Define in C the following curve:

$$A(\lambda) = \bar{f}_\eta^\lambda(A) = f_\eta^\lambda{}^*(A)$$

for some given $A \in C$. The tangent vector of this curve in A :

$$\frac{d}{d\lambda}A(\lambda) = \frac{d}{d\lambda}[\exp(\lambda\eta)]^*A\Big|_{\lambda=0}$$

turns out to be (see Bleeker, Theorem 3.2.16, [16]):

$$\frac{d}{d\lambda}A(\lambda) = D_A\eta \ . \qquad (3.6.5)$$

One can think about the tangent vector to the curve A( ) as the infinitesimal generator in C of the "gauge" transformation, induced by η .

Notice that the vector (3.6.5) is just the $\frac{\delta}{\delta A}$ component of the gauge vector (3.6.4). This is the motivation to lift the action of $\overline{GA(P)}$ on C to an action on the cotangent bundle T^*C in the obvious way, i.e. by the pullback of $\bar{f}$: we define a transformation $T^*\bar{f}$,

$$T^*\bar{f} : T^*C \to T^*C$$

by

$$T^*f(A,\pi) = (\bar{f}^{-1}(A), \bar{f}^*(\pi)) \ .$$

See figure 6.

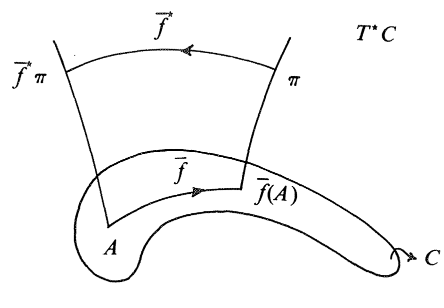


Fig. 6. The lift of a diffeomorphism of C to the cotangent bundle T^*C

(In classical mechanics this procedure of lifting transformations from configuration space to phase space is well-known. One speaks of point transformations. See Abraham & Marsden [2], p.181).

We now determine the form of an infinitesimal transformation on T^*C induced by η, i.e. we calculate the tangent to the curve

$$(A(\lambda),\pi(\lambda)) = T^*\bar{f}^{\lambda}_{\eta}(A,\pi) .$$

THEOREM 3.6.1.

$$\frac{d}{d\lambda}(A(\lambda),\pi(\lambda))\Big|_{\lambda=0} = -D_A\eta\frac{\delta}{\delta A} - [\pi,\eta]\frac{\delta}{\delta\pi} .$$

PROOF

To first order in λ we have, using (3.6.5),

$$A(\lambda) = \left(\bar{f}^{\lambda}_{\eta}\right)^{-1}(A) = A - \lambda D_A\eta .$$

This means that $(A(\lambda);\pi(\lambda))$ a 1-form is in the point $A - \lambda D_A\eta$. To calculate $\pi(\lambda)$ we contract this form with a vector $X = (A(\lambda),\dot{A})$:

$$<(A(\lambda),\pi(\lambda)) \mid (A(\lambda),\dot{A})> = <\pi(\lambda) \mid \dot{A}> .$$

With the definition of $\pi(\lambda)$ we have

$$<\pi(\lambda) \mid \dot{A}> = <\pi \mid \bar{f}^{\lambda}_{\eta*}\dot{A}> . \tag{3.6.6}$$

We represent X by a curve through $A(\lambda)$:

$$A(\lambda,\tau) = A(\lambda) + \tau\dot{A}$$

(with $\frac{d}{d\tau}A(\lambda,\tau) = \dot{A}$) .

Then

$$\bar{f}^{\lambda}_{\eta*}\dot{A} = \bar{f}^{\lambda}_{\eta*}\frac{d}{d\tau}A(\lambda,\tau) = \frac{d}{d}\,\bar{f}^{\lambda}_{\eta}(A(\lambda,\tau))$$

$$= \frac{d}{d\tau}(A(\lambda,\tau) + \lambda D_{A(\lambda,\tau)}\eta) =$$

$$= \frac{d}{d\tau}(A(\lambda) + \tau\dot{A} + \lambda(d\eta + [A(\lambda) + \tau\dot{A},\eta]))$$

$$= \dot{A} + \lambda[\dot{A},\eta] .$$

Substituting in (3.6.6):

$$<\pi(\lambda) \mid \dot{A}> = <\pi \mid \dot{A} + \lambda[\dot{A},\eta]>$$

$$= <\pi \mid \dot{A}> + \lambda<\pi \mid [\dot{A},\eta]> .$$

The contraction $<|>$ is defined using the trace of the Lie algebra. Using the cyclicity of the trace we find

$$<\pi(\lambda) \mid \dot{A}> = <\pi - \lambda[\pi,\eta] \mid \dot{A}> .$$

Hence

$$(A(\lambda), \pi(\lambda)) = (A-\lambda D_A\eta, \pi - \lambda[\pi,\eta]) .$$

The tangent to this curve is

$$-D_A\eta \frac{\delta}{\delta A} - [\pi,\eta]\frac{\delta}{\delta\pi} . \quad ∎$$

Comparing this result with (3.6.4) we see that the set of gauge vectors of the Yang-Mills system (after eliminating the A_o component) coincides with the infinitesimal generators of "gauge" transformations on T^*C. (The minus sign is irrelevant since both η and $-\eta$ generate gauge transforms (and also "gauge" transformations!). The minus sign comes from the definition of the action of GA(P) on T^*C by pullback, see also fig. 6).

Notice that we have proved in section 3.3 that gauge vectors are tangent to $\bar{M}_2$, so "gauge" transformations map points of $\bar{M}_2$ to points of $\bar{M}_2$.

Now we want to understand the meaning of the final constraint manifold $\bar{M}_2$ and of the condition

$$\delta_A \pi = 0$$

from the point of view of "gauge" transformations. In the above we have seen that any element η of $\bar{\Lambda}^o(P_3, \ell(G))$ generates "gauge" transformations on T^*C and has associated with it a vector field, its infinitesimal generator:

$$X_\eta = D_A\eta \frac{\delta}{\delta A} + [\pi,\eta] \frac{\delta}{\delta \pi} \quad . \tag{3.6.7}$$

T^*C , as a cotangent bundle, is a symplectic manifold with a canonical 2-form ω derived from the Liouville-form θ by $\omega = d\theta$. One might wonder whether the gauge vector field (3.6.7) is the Hamiltonian vector field of some Hamiltonian function. This turns out to be the case:

<u>THEOREM 3.6.1</u>.

$$i(X_\eta)\omega = -dH_\eta$$

with

$$H_\eta = (\pi; D_A\eta) \ .$$

<u>PROOF</u>.

One can prove that the Liouville-form θ is invariant under "gauge" transformations:

$$(T^*\bar{f}_\eta)^*\theta = \theta \ .$$

(In fact θ is invariant under the lift of <u>any</u> diffeomorphism of the base space, see Abraham & Marsden [2], Theorem 3.2.12). Since X_η is the infinitesimal generator of $T^*\bar{f}_\eta$, we find that the Lie derivative of θ in the direction X_η is zero:

$$L_{X_\eta}\theta = \frac{d}{d\lambda}\left[\frac{(T^*\bar{f}_\eta^\lambda)^*\theta-\theta}{\lambda}\right] = 0 \ .$$

Using property A.8 of the Lie derivative and the definition of ω we obtain:

$$L_{X_\eta}\theta = di(X_\eta)\theta + i(X_\eta)d\theta = 0$$

or

$$i(X_\eta)\omega = -d[i(X_\eta)\theta] \ .$$

Hence

$$H_\eta = i(X_\eta)\theta \ .$$

Writing θ in local coordinates and using the form (3.6.7) for X_η we find

$$H_\eta = \langle\theta \mid X_\eta\rangle = \langle\pi\delta A \mid D_A\eta\frac{\delta}{\delta A} + [\pi,\eta]\frac{\delta}{\delta\pi}\rangle$$

$$= (\pi, D_A\eta) \ . \quad \blacksquare$$

Notice that the mapping

$$\eta \to H_\eta$$

from the Lie algebra $\bar{\Lambda}^o(P_3,\ell(G))$ of the "gauge transformations to Hamiltonian functions is linear. (It is in fact a Lie algebra homomorphism if we define the Lie product of two functions H_η and $H_{\eta'}$ by the Poisson-bracket:

$$\left\{H_\eta, H_{\eta'}\right\} = \omega\left(X_{H_\eta}, X_{H_{\eta'}}\right)$$

Because of this linearity we can define another function:

$$J: T^*C \to \bar{\Lambda}^o(P_3,\ell(G))^*$$

which maps points of T^*C to the dual of the Lie algebra of "gauge" transformations, such that

$$H_\eta(x) = \langle J(x) \mid \eta \rangle \qquad \forall x \in T^*C .$$

With $\langle | \rangle$ the contraction between the "gauge" Lie algebra and its dual. This J is Souriau's *momentum mapping*, (see Souriau [7], Abraham & Marsden [2]). In our case it is:

$$J(A,\pi) = \delta_A \pi .$$

We find that the final constraint manifold $\bar{M}_2$ is the inverse image of zero of the momentum mapping:

$$\bar{M}_2 = J^{-1}(0) .$$

This is quite a remarkable result: $\bar{M}_2$, the final constraint manifold of a presymplectic system (M_1, ω_1, H) appears here as a level manifold of a function J which has nothing to do with the precise dynamics of the system we are studying. (J does not depend, in any way, on the form of the Hamiltonian (3.6.1)). J depends only on the underlying geometry of the Yang-Mills system.

We proceed to calculate the reduced phase space associated to $\bar{M}_2$ by quotienting $\bar{M}_2$ by the action of the lift of $\overline{GA(P)}$ to T^*C . We denote this lift by $GA(P)$.

<u>THEOREM 3.6.2.</u>

$$\bar{M}_{2/GA(P)} \cong T^*(C/\overline{GA(P)}) = T^*W$$

with $\quad W = C/\overline{GA(P)}$

PROOF.

We have to construct two smooth mappings between T^*W and $\bar{M}_{2/GA(P)}$, which are each other's inverse. First we will show that any form in T^*W can be pulled back by the projection from C to W to an equivalence class of forms on $\bar{M}_2$, (Lemma 1, 2 and 3). This gives the mapping $T^*W \to \bar{M}_{2/GA(P)}$. Then we define a mapping from $\bar{M}_2$ to T^*W and prove that equivalent points of $\bar{M}_2$ map on the same point of T^*W. This gives the mapping $\bar{M}_{2/GA(P)} \to T^*W$.

We need some definitions: let

$$pr: C \to W$$

be the projection which associates to a connection A its equivalence class $[A]$ in W. The pullback for pr gives a mapping

$$pr^*: T^*W \to T^*C$$

Let $w \in T^*_{[\bar{A}]}W$, with $[\bar{A}]$ some fixed equivalence class, then pr^*w is a collection of forms in T^*C, all lying above points A that project on fixed equivalence class $[\bar{A}]$:

$$pr^*w = \{(A,\pi) \in T^*C \mid pr(A) = [\bar{A}] \ , \ \pi(X) = w(pr_*X), \ \forall X \in T_AC\}$$

LEMMA 1.

Any $(A,\pi) \in pr^*w$ belongs to $\bar{M}_2$, for all $w \in T^*W$.

LEMMA 2.

All points of pr^*w are related through transformations of $GA(P)$, for all $w \in T^*W$.

LEMMA 3.

If $(A,\pi) \in pr^*w$, then also $T^*\bar{f}(A,\pi) \in pr^*w$, for all $f \in GA(P)$ and all $w \in T^*W$.

PROOF of lemma 1.

Let $X \in T_A C$ be a vector tangent to an orbit of $GA(P)$ in C , then the projection $pr_* X$ of X to $T_{[A]}W$ is zero:

$$pr_* X = 0 .$$

Consequently

$$pr^* w(X) = w(pr_* X) = 0 .$$

Such an X is of the form (see equation (3.6.5))

$$X = D_A \eta \frac{\delta}{\delta A} .$$

Let (A,π) be a point of pr^*w , then

$$< \pi \delta A \mid D_A \eta \frac{\delta}{\delta A} > = (\pi ; D_A \eta) = 0 , \quad \forall \eta$$

and (A,π) belongs to $\bar{M}_2$. ▼

PROOF of lemma 2.

Let (A,π) and (A',π') in pr^*w . A and A' project on the same class $[\bar{A}]$, hence

$$A' = \bar{f}(A)$$

for some f in $GA(P)$. We have

$$\pi(X) = w(pr_* X) , \quad \forall X \in T_A C$$

but

$$pr = pr \circ \bar{f}$$

(because $\bar{f}$ is a "gauge" transformation). Therefore

$$\pi(X) = w(pr \circ \bar{f})_* X) = w(pr_* \bar{f}_* X) = pr^* w(\bar{f}_* X)$$

$$= \pi'(f_* X) = f^* \pi'(X)$$

and we find

$$(A,\pi) = (\bar{f}^{-1}(A'), f^*\pi') = T^*\bar{f}(A',\pi')$$

and (A,π) and (A',π') are gauge related. ▼

PROOF of lemma 3.

Let $(A,\pi) \in pr^*w$ and $(A',\pi') = T^*\bar{f}(A,\pi) = (\bar{f}^{-1}(A), \bar{f}^*(\pi))$. We have to show that $(A',\pi') \in pr^*w$. This is the case if

1) $pr(A') = [A] = pr(A)$

but $A' = \bar{f}^{-1}(A)$ and $pr \circ \bar{f}^{-1} = pr$, so $pr(A') = pr(A)$

2) $\pi'(X) = w(pr_* w)$

but $\pi'(X) = \bar{f}^*\pi(X) = \pi(\bar{f}_* X) = w(pr_* \bar{f}_* X)$

$= w(pr \circ f)_* X) = w(pr_* X)$. ▼

These three lemmata show that any element of T^*W lifts by pullback to an equivalence class of points of $\bar{M}_2$, i.e. to an element of $\bar{M}_{2/GA(P)}$. We show conversely that every equivalence class of $\bar{M}_2$ defines a unique element of T^*W. To this end define a mapping

$$s : \bar{M}_2 \to T^*W$$
$$(A,\pi) \to w \in T^*_{[A]}W$$

by

$$w(X) = \pi(\tilde{X})$$

with

$$X \in T_{[A]}W \; , \quad \tilde{X} \in T_A C \; , \quad X = pr_* \tilde{X}$$

<u>LEMMA 4</u>.

The definition of s is independent of the choice of vector $\tilde{X}$ that projects on X.

<u>LEMMA 5</u>.

Points of $\bar{M}_2$ that are related through "gauge" transformations have the same image under s.

<u>PROOF of lemma 4</u>.

Let $\tilde{X}_1$ and $\tilde{X}_2$ both project on X. This means that they differ by a gauge vector:

$$\tilde{X}_1 = \tilde{X}_2 + D_A \eta$$

then

$$w(X) = \pi(\tilde{X}_1) = \pi(\tilde{X}_2) + \pi(D_A \eta)$$

but the last term:

$$\pi(D_A \eta) = (\pi ; D_A \eta)$$

is zero because we are on $\bar{M}_2$. ▼

PROOF of lemma 5.

Let $(A',\pi') = T^*\bar{f}(A,\pi)$, $w = s(A,\pi)$ then

$$\pi'(X) = \pi(\bar{f}_*\tilde{X} = w(pr_* \circ \bar{f}_*\tilde{X}) = w(pr_*X) = \pi(X)$$

and accordingly

$$s(A',\pi') = w . \quad \blacktriangledown$$

The mapping s is the inverse of the lifting by pr^* :

$$s(pr^*w) = w ,$$

and also

$$pr^*(s((A,\pi))) = (A,\pi) .$$

We leave it to the diligent reader to prove that the mappings between T^*W and $\bar{M}_{2/GA(P)}$ are smooth and that therefore these spaces are diffeomorphic. ∎

REMARK.

$\bar{M}_{2/GA(P)}$ as a reduced phase space has a symplectic structure defined on it (see section 2.6). T^*W has a canonical symplectic structure. These two structures are identical if we make the identification implied by the last theorem.

Theorem 3.6.2 tells us that the reduced phase space for the Yang-Mills theory can be obtained by first eliminating the unphysical (and ungeometrical) degrees of freedom from the configuration space C to get W , the "true" configuration space and then constructing T^*W , the true phase space.

The reduction of the space of connections C by the action of the group of gauge transformations GA(P) has been studied extensively in recent years (Singer [19], Narashiman and Ramadas [29], Mitter and Viallet [22]). We will summarize the results of these investigations, referring for proofs more to the literature than we have done until now, since one needs for those somewhat more advanced mathematics (cohomology etc.) than we have used in this report.

To perform the reduction of C by the action of $GA(P)$ we have to impose some restrictions on both C and $GA(P)$.

Notice first that a gauge transformation

$$f_z : P \to P \ , \qquad f_z(p) = pz = R_z p$$

with z a constant element of the centre of G (so $zg = gz$ for all $g \in G$) has trivial action on C :

$$\bar{f}_z(A) = f_z^* A = R_z^* A \overset{(1)}{=\!=} Ad_z A \overset{(2)}{=\!=} (i_z)_* A \overset{(3)}{=\!=} (id)_* A = A \ .$$

For equality (1) we use the fact that a connection transforms with the adjoint representation, (2) is the definition of the adjoint representation (with $i_z g = zgz^{-1}$) and (3) uses the fact that z lies in the centre of G, so $i_z g = zgz^{-1} = gzz^{-1} = id\, g$ for all $g \in G$. To get a not trivial group action on C we have to qoutient $GA(P)$ by the constantly centre valued transformations. We denote the resulting group by $\overline{GA(P)}$.

Next consider a connection A in C for which

$$A = A + D_A \eta$$

for some generator of gauge transformations η . Such a connection A , which admits one or more covariantly constant functions η is called reducible. One can prove (Kobayashi and Nomizu [38]) that a principle fiber bundle $P(M,G)$ with a reducible connection is equivalent to a reduced bundle

P(M,H) with H a subgroup of G, with the same connection A.
The "dimension" of the orbit of GA(P) through a reducible connection is smaller than the orbit through an irreducible connection (one that does not admit covariantly constant functions). One can not properly speak of "dimension" in the infinite dimensional spaces we are working with, but the intuitive idea will be clear: some generators η leave invariant reducible connections while transforming irreducible ones. This "jump" in "dimension" gives rise to singularities, which are studied by Arms [33]. To avoid these singularities one restricts the action of $\overline{GA(P)}$ to the irreducible connections, denoted by $\bar{C}$. One can prove [19] that $\bar{C}$ is open and dense in C, so by this restriction to $\bar{C}$ one does not loose much information. One then proves ([19], [29], [22]) that (modulo technical details)

$$\pi: \bar{C} \to \bar{W} = \bar{C}/\overline{GA(P)} \tag{3.6.8}$$

is a C^{∞}-principle fiber bundle with $\bar{W}$ a smooth manifold. $\bar{W}$ is called the orbit space of Yang-Mills theory.

The Riemannian structure of $\bar{C}$ enables us to define a connection in the fibration (3.6.8). We take as horizontal vectors in a point A of $\bar{C}$ those vectors which are orthogonal to the fiber through A. Since the tangents to the fiber are of the form $D_A\eta$, horizontal vectors τ satisfy:

$$(D_A\eta;\tau) = 0$$

or

$$\delta_A\tau = 0 .$$

The connection form associated to this horizontality is

$$\chi_A = (\delta_A D_A)^{-1}\delta_A$$

as one can easily check.

The base space $\bar{W}$ inherits from $\bar{C}$ a Riemannian structure: Let $w \in \bar{W}$, and $X_w, Y_w \in T_w\bar{W}$

$$X_w = \pi_* X_A, \quad Y_w = \pi_* Y_A, \quad \text{for some } X_A, Y_A \in T_A\bar{C}, \quad \pi(A) = w .$$

Then we define the metric $g_{\bar{W}}$ on $\bar{W}$ by

$$g_{\bar{W}}(X_w, Y_w) = (X_A; Y_A)$$

with $(\, ; \,)$ the inner product of $T_A\bar{C} = \bar{\Lambda}^1(P_3, \ell(G))$.

The Hamiltonian $H(A,\pi)$ on $\bar{M}_2$ is invariant under gauge transformations and therefore leads to a Hamiltonian $\tilde{H}$ on $T_*\bar{W}$, which is defined by

$$\bar{H}(w, \pi_w) = H(A,\pi) = -\tfrac{1}{2}\rho(\pi;\pi) + \tfrac{1}{2}(D_A A; D_A A)$$

with (A,π) any point on $\bar{M}_2$ that projects on (w,π_w) . We see that in the Minkowski case $(\rho = -1)$ the Hamiltonian on $T^*\bar{W}$ is of the familiar form: a kinetic part quadratic in the momenta and a potential $V(A) = \tfrac{1}{2}(D_A A; D_A A)$ independent of the momenta. Babelon and Viallet [23] have studied this system with the potential $V(A)$ put equal to zero. The projections (from T^*W to W) of integral curves of solutions of the equations of motion are then geodesics with respect to the metric $g_{\bar{W}}$ defined above. They discovered an infinity of conserved quantities that are in involution, suggesting that the geodesic motion on $\bar{W}$ is completely integrable. It would be interesting to investigate the influence of the potential $V(A)$ on these results.

One can introduce explicit coordinates on $\bar{W}$. Take a point $w \in \bar{W}$ and a point A in $\bar{C}$ projecting on it: $\pi(A) = w$. Define a linear subspace S_A of $\bar{C}$ by

$$S_A = \{A + \tau \in \bar{C} \mid \delta_A \tau = 0\} .$$

This submanifold S_A intersects, as we have seen, the fiber through A orthogonal, and by continuity it will intersect the fibers in a neighbourhood of A only one time. This means that S_A defines a (local) section of the fibration (3.6.8). We use S_A to coordinatize $\bar{W}$. Let w' be a point in a neighbourhood U of w, such that there exists a unique $A+\tau$ in S_A that projects on w'. We take as a coordinate function in U

$$\phi : U \subset \bar{W} \to \bar{\Lambda}^1(P_3, \ell(G))$$

$$w' = \pi(A+\tau) \to \tau .$$

In the physics literature the coordinatization of $\bar{W}$ by choosing a section of (3.6.8) is called fixing a gauge. One refers to the particular coordinatization by ϕ defined above as the background gauge or the generalized Coulomb gauge.

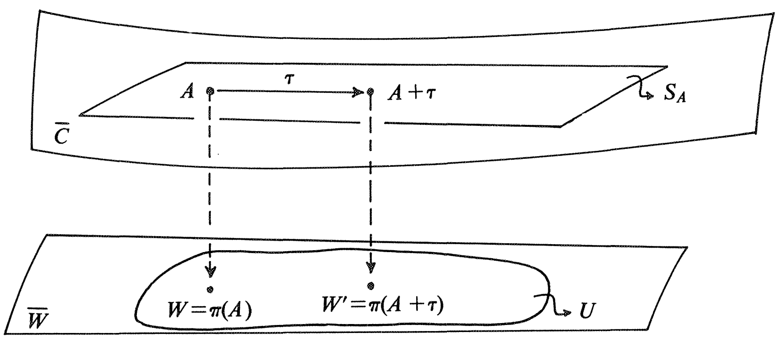


Fig. 7. The background gauge.

This coordinate system is used in the path-integral method of quantization. Formally one should integrate over all paths in $\bar{W}$ to obtain the generating functional:

$$Z = \int_{\bar{W}} \mathcal{D}w \, e^{-iS(w)/\hbar}$$

Using, however, the diffeomorphism

$$\pi_A : S_A \to \bar{W}$$

with π_A the restriction of $\pi : \bar{C} \to \bar{W}$ to S_A, one can equally well-integrate over S_A:

$$\begin{aligned} Z &= \int_{\pi_A(S_A)} \mathcal{D}w\, e^{-iS(w)/\hbar} \\ &= \int_{S_A} \pi_A^* \mathcal{D}w\, e^{-iS(\pi_A^* w)/\hbar} \\ &= \int_{S_A} \det(\pi_A) \mathcal{D}\tau\, e^{-iS(\tau)/\hbar} \end{aligned}$$

$\mathrm{Det}(\pi_A)$ is the Jacobian determinant of the mapping π_A.

It is the famous Faddeev-Popov determinant. One can also interpret it as the square root of the Riemannian metric $g_{\bar{W}}$, as was shown by Babelon and Viallet [40].

We have in this (admittedly sketchy) discussion of the path integral assumed that S_A intersects every fiber exactly once, in other words, we assumed that S_A defines a global section of the fibration (3.6.8). Gribov [39] has discovered that this assumption is incorrect if one studies non-abelian theories in $\mathbb{R}^4$ with boundary conditions equivalent to compactifying space to S^3. He showed that S_A far away from A intersects in this situation the orbit through A again, see fig. 8.

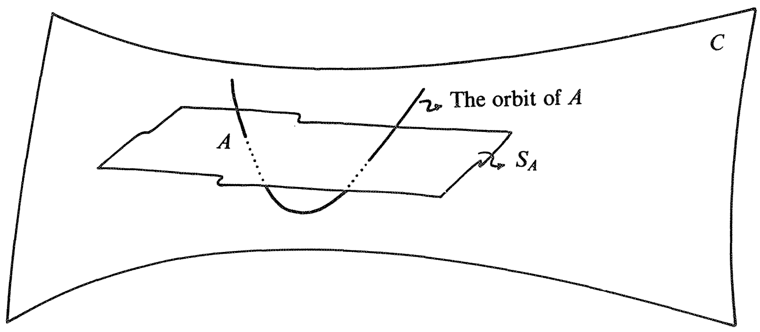


Fig. 8. The Gribov effect.

When this occurs S_A fails to be a global section of the fibration (3.6.8). One might then wonder whether there exist any global sections of this fibration at all. Singer [19] and independently Narashiman and Ramadas [20] have proved that if one takes space to be compact and G non-abelian there exist indeed <u>no</u> global sections of the fibration (3.6.8). The fibration is nontrivial. This has the consequence that one needs more than one coordinate neighbourhood (in fact a countable infinite number, Singer [37]) to coordinatize $\bar{W}$ by local sections of (3.6.8). This means that the path-integral method as described above is strictly not correct. In some situations, however, one can treat the theory pertubatively, i.e. consider only small variations $\tilde{A}$ around some fixed connection A. Then the above method will lead to the familiar Feynman graphs. To include nonpertubative effects, i.e. the influence of large fluctuations, one should calculate the path-integral in local sections of (3.6.8) above neighbourhoods in $\bar{W}$ forming a partition of unity and then sum over all neighbourhoods. For this to be practical one should have explicit coordinates of $\bar{W}$ independent of sections of (3.6.8).

SUMMARY AND OUTLOOK

We have now reached the goal of this report: to derive the true phase space of Yang-Mills theory. The true phase space turned out to be the cotangent bundle of orbit space, the space of the gauge inequivalent connections in $P_3(M_3,G)$. Of course, this is only kinematics and there remains a lot to be done before one can say that one understands the dynamics of the Yang-Mills theory, even at a classical level, not to mention the quantized theory.

For instance, one would like to know how the self-dual solutions of Yang-Mills equations (those for which the curvature (in not 3+1 decomposed form) satisfies $*F(A) = \pm F(A)$) are embedded in orbit space. These self-dual solutions have interesting properties (Bäcklund transformations, infinite number of conserved quantities) suggesting complete integrability and one wonders whether these properties might extend to the complete Yang-Mills theory. One way of investigating this question could be to translate these results in the symplectic language described in this report. Research along these lines is in progress.

Appendix A.

In this appendix we prove the following

THEOREM.

Let (P,ω,H) be a presymplectic system, M_k its final constraint manifold, X_H a solution tangent to M_k of the equations of motion:

$$i(X_H)\omega/M_k = -dH/M_k \tag{A.1}$$

and let

$$G_\ell = G_{\ell-1} + [G_{\ell-1}, G_{\ell-1}] + [G_{\ell-1}, X_H] \qquad \ell = 2,3,\ldots$$

with

$$G_1 = \ker\Omega \cap \underline{TM}_k\ .$$

Then

$$G_\ell \subset TM_k^\perp \cap \underline{TM}_k \tag{A.2}$$

PROOF.

We use induction on ℓ. Let ℓ be equal to 1. If $Y \in G_1$ then $Y \in \ker\Omega$, i.e.

$$\omega(Y,Z) = 0$$

for all Z, in particular for Z of the form

$$Z = j_*\tilde{Z}$$

(with $\tilde{Z} \in TM_k$ and $j: M_k \to P$ the embedding of M_k in P). We find therefore

$$Y \in TM_k^\perp$$

Y is also an element of $\underline{TM}_k$ and the theorem is proved for $\ell = 1$.

Let (A.2) be true for some ℓ_o :

$$G_{\ell_o} \subset TM_k^{\perp} \cap \underline{TM}_k .$$

This means two things

1. $j^* i(G_{\ell_o}) = 0$

2. G_{ℓ_o} is tangent to M_k .

We have to prove that 1. en 2. are both true for $\ell_o + 1$:

1. $j^* i(G_{\ell_o+1}) = 0$ (A.3)

2. G_{ℓ_o+1} is tangent to M_k (A.4)

with

$$G_{\ell_o+1} = G_{\ell_o} + \left[X_H, G_{\ell_o}\right] + \left[G_{\ell_o}, G_{\ell_o}\right] \tag{A.5}$$

That (A.4) is true is obvious: G_{ℓ_o} is tangent to M_k by the induction hypothesis, X_H is also tangent to M_k and sums and commutators of vectors tangent to a submanifold are again tangent to that submanifold.

To prove (A.3) we have to make a little compoutation. Substitution of (A.5) in (A.3) gives an expression with three terms:

$$j^* i(G_{\ell_o+1})\omega = j^* i(G_{\ell_o})\omega + j^* i(\left[G_{\ell_o}, G_{\ell_o}\right])\omega + \\ + j^* i(\left[X_H, G_{\ell_o}\right])\omega \tag{A.6}$$

The first term is zero by the induction hypothesis.

To calculate the other terms we use two identities:

$$i([X,Y]) = L_X i(Y) - i(Y) L_X \tag{A.7}$$

$$L_X = di(X) + i(X)d \tag{A.8}$$

with L_X the Lie derivative in the direction X, see for example Chapter 2 of Abraham and Marsden [2].

Using this, the second term of (A.6) becomes:

$$\begin{aligned} j^*i([Z_1, Z_2])\omega &= j^*(\left\{L_{Z_1} i(Z_2) - i(Z_2)L_{Z_1}\right\}\omega) = \\ &= j^*(\{di(Z_1)i(Z_2) + i(Z_1)di(Z_2) - i(Z_2)di(Z_1) - i(Z_2)i(Z_1)d\}\omega) \end{aligned} \tag{A.9}$$

Z_1 and Z_2 are arbitrary vectors in G_{ℓ_o}. The last term of (A.9) is zero because ω is closed. The other terms are zero because Z_1 and Z_2 both belong to $\underline{TM}_k$ and to $TM_k^{\perp}$ by the induction hypothesis. As an example we will show that the second term of (A.9) is zero:

$$j^*[i(Z_1)di(Z_2)\omega](\tilde{Y}) = 0 \tag{A.10}$$

for arbitrary $\tilde{Y}$ in TM_k. For convenience we introduce the 1-form α by:

$$\alpha = i(Z_2)\omega .$$

Remark that α is zero on elements of TM_k (since $Z_2 \in TM_k^{\perp}$). Define the image of $\tilde{Y}$ in $\underline{TM}_k$ by:

$$Y = j_*\tilde{Y} .$$

With this notation the left-hand side of (A.10) becomes

$$\begin{aligned} j^*i(Z_1)di(Z_2)\omega(\tilde{Y}) &= i(Z_1)d\alpha(Y) \\ &= d\alpha(Z_1,Y) . \end{aligned}$$

Using the definition of exterior derivative we get:

$$d\alpha(Z_1,Y) = Z_1\alpha(Y) - Y\alpha(Z_1) - \alpha([Z_1,Y])$$

and this is indeed zero because Y, Z_1 and $[Z_1,Y]$ all belong to $\underline{TM}_k$ and, as we have already said, α annihilates $\underline{TM}_k$. Similar calculations show that the other terms of (A.9) are also zero and so we find that the second term of (A.6) does not contribute.

We now calculate the last term of (A.6), again using (A.7) and (A.8):

$$\begin{aligned} j^*i([Z,X_H])\omega &= j^*[\{L_Z i(X_H) - i(X_H)L_Z\}\omega] \\ &= j^*[\{di(Z)i(X_H) + i(Z)di(X_H) - i(X_H)di(Z) - i(X_H)i(Z)d\}\omega] \, . \end{aligned} \tag{A.11}$$

The fourth term of (A.11) is zero because ω is closed, the first and third terms can be shown to be zero by the same type of calculation as used to prove (A.10). So we find for (A.6):

$$j^*\Big[i(G_{\ell_o+1})\omega\Big] = j^*\Big[i(G_{\ell_o})di(X_H)\omega\Big] \tag{A.12}$$

Using the equations of motion (A.1) we can write this as

$$j^*\Big[i(G_{\ell_o+1})\omega\Big] = j^*\Big[i(G_{\ell_o})d(dH/M_k)\Big]$$

Naively one could think that

$$d(dH/M_k) = 0$$

but things are a little bit more complicated than that. The problem is that by restricting dH to M_k one can loose the exactness of it: dH/M_k is an element of T^*P/M_k and one has to extend this form to T^*P to calculate the exterior derivative of it. This extension is in general not equal to dH.

To be more explicit we introduce submanifold charts on P, i.e. charts (U,ϕ) such that

$$\phi: U \subset P \to E \times F$$

and

$$\phi(U \cap M_k) = \phi(U) \cap E\times\{0\}$$

In such a chart a point of U has coordinates (p_E, p_F), with $p_E \in E$, $p_F \in F$ and if it has coordinates of the form (p_E, o) it lies on M_k.

For simplicity we assume both E and F to be finite dimensional. Introduce bases $\{e_i\}$ and $\{f_j\}$ for E and F respectively, and their dual bases $\{de_i\}$ and $\{df_i\}$, with

$$de_i(e_j) = \delta_{ij}, \quad df_i(f_j) = \delta_{ij}, \quad de_i(f_j) = 0 .$$

With respect to these bases we have

$$p = (p_E, p_F) = p_E^i e_i + p_F^i f_i$$
$$H(p) = H(p_E^i e_i, p_F^i f_i)$$

and

$$dH(p) = \frac{\partial H}{\partial e_i}(p_E^i, p_F^i) de^i + \frac{\partial H}{\partial f_j}(p_E^i, p_F^j) df^j .$$

Restricting dH to M_k we get:

$$dH/M_k = \frac{\partial H}{\partial e_i}(p_E^i) de^i + \frac{\partial H}{\partial f_j}(p_E^i) df^j$$

where the partial derivatives do not any longer depend on the coordinates in F.

To calculate the exterior derivative we extend dH/M_k constantly in the direction of F, we find

$$d(dH/M_k) = \frac{\partial^2 H}{\partial e_k \partial e_i} de^k \wedge de^i + \frac{\partial^2 H}{\partial e_k \partial f_i} de^k \wedge df^i .$$

The first term is zero because of the symmetry of the partial derivatives and the antisymmetry of the wedge product. So we end up with

$$d(dH/M_k) = \frac{\partial^2 H}{\partial e_k \partial f_i} de^k \wedge df^i$$

To determine (A.13) we have to calculate:

$$i(Z)d(dH/M_k) = \frac{\partial^2 H}{\partial e_k \partial f_i}\left[de^k(Z)df^i - de^k df^i(Z)\right].$$

Now Z belongs to G_{ℓ_o}, therefore to TM_k and so Z does not have components in the direction of F:

$$df^i(Z) = 0 \qquad \forall i .$$

Substituting all this in (A.13) we find

$$j^*\left[i(G_{\ell_o+1})\omega\right] = j^*\left(\frac{\partial^2 H}{\partial e_k \partial f_i} de^k(Z)df^i\right) = 0$$

Because the pullback of df^i to M_k is zero:

$$j^*df^i = dj^+f^i = 0 .$$

We leave it to the reader to check that our final result is independent of the choice of coordinates we have made, and to extend the proof to infinite dimensional spaces.

REFERENCES

[1] SUDARSHAN, E.L.G. - MUKUNDA, N.: *Classical Dynamics: A modern Perspective*, Wiley, New York, 1974.

[2] ABRAHAM, R. - MARSDEN, J.E.: *Foundations of Mechanics*, second edition, Benjamin, Reading (Mass.), 1978.

[3] ARNOLD, V.I.: *Mathematical Methods of Classical Mechanics*, Springer, New York, 1978.

[4] GIACHETTI, R.: Rev. Nuovo Cimento 4(1981) 1-63, "Hamiltonian systems with symmetry: an introduction".

[5] CHERNOFF, P.R. - MARSDEN, J.E.: *Properties of infinite dimensional Hamiltonian Systems*, Springer, Berlin, 1974.

[6] MARSDEN, J.E.: *Applications of Global Analysis in Mathematical Physics*, Publish or Perish, Boston, 1974.

[7] SOURIAU, J.M.: *Structure des Systèmes Dynamiques*, Dunod, Paris, 1970.

[8] DIRAC, P.A.M.: *Lectures on Quantum Mechanics*, Yeshiva University, New York, 1964.

[9] GOTAY, M.J. - NESTER, J.M. - HINDS, G.: J.Math.Phys. 19(1978) 2388-2399: "Presymplectic manifolds and the Dirac-Bergmann theory of constraints".

[10] LICHNEROWITZ, A.: C.R.Acad.Sci. Paris A280(1975) 523-527: "Variété symplectique et dynamique associé a une sous-variété".

[11] GOTAY, M.J. - NESTER, J.M.: "Presymplectic Hamilton and Lagrange systems, gauge transformations and the Dirac theory of constraints" in: *Grouptheoretical Methods in Physics*, W. Beigelbock(ed.), Springer, Berlin, 1979.

[12] GOTAY, M.J.: "On the validity of Dirac's conjecture regarding first class constraints: Preprint, University of Calgary.

[13] SNIATYCKI, J.: Ann.Inst. H. Poincaré, Sect.A, XX(1974) 365-372: "Dirac brackets in geometric dynamics".

[14] MacFARLANE, A.J.: "Equations of Korteweg-de Vries type I: Lagrangian and Hamilton formalism". Preprint CERN TH 3289, 1982.

[15] SUNDERMEYER, K.: *Constrained Dynamics*, Springer, Berlin, 1982.

[16] BLEEKER, D.: *Gauge theory and variational principles*, Addison-Wiley, Reading (Mass.), 1981.

[17] EGUCHI, T. - GILKEY, P.B. - HANSON, A.J.: Phys.Rep. 66(1980) 213-393, "Gravitation, gauge theories and differential geometry".

[18] DANIEL, M. - VIALLET, C.M.: Rev.Mod.Phys. 52(1980) 174-197, "The geometrical setting of gauge theories of the Yang-Mills type".

[19] SINGER, I.M.: Commun.Math.Phys. 60(1978) 7-12: "Some remarks on the Gribov ambiguity".

[20] NARASHIMHAN, M.S. - RAMADAS, T.R.: Commun.Math.Phys. 67(1979) 121-136: "Geometry of SU(2) Yang-Mills Fields".

[21] BALELON, O. - VIALLET, C.M.: Commun.Math.Phys. 81(1981) 515-525, "The Riemannian geometry of the orbit space of gauge theories".

[22] MITTER, P.K. - VIALLET, C.M.: Commun.Math.Phys. 79(1981) 457-472, "On the bundle of connections and the gauge orbit manifold in Yang-Mills theory".

[23] BABELON, O. - VIALLET, C.M.: Phys.Lett. 103B(1981) 45-47: "Conserved quantities for the geodesic motion on the configuration space of non-abelian Yang-Mills theory".

[24] MITTER, P.K.: "Geometry of the space of gauge orbits and the Yang-Mills dynamical system", in: *Recent developments in gauge theories*, G. 't Hooft(editor), Cargèse Lectures (1979), Plenum, New York, 1979.

[25] MISNER, C.W. - THORN, K.S. - WHEELER, J.A.: *Gravitation*, Freeman, San Francisco, 1970.

[26] BUUR, J.N.: *Differential geometry and classical field theory*, Doctoraal scriptie, Instituut voor Theoretische Fysica, Amsterdam, 1982.

[27] BACHMAN, G. - NARICI, N.: *Functional Analysis*, Academic Press, New York, 1966.

[28] BERGMANN, P.G.: Helv.Phys.Acta.suppl. IV(1956) 79-95: "Quantisering allgemein-kovariante Feldtheorien".

[29] BISHOP, R.L. - GRITTENDEN, R.J.: *Geometry of Manifolds*, Academic Press, New York, 1964.

[30] WEINSTEIN, A.: *Lectures on symplectic manifolds*, American Mathematical Society, Providence (R.I), 1977.

[31] SCHUTZ, B.: *Geometrical methods of Mathematical Physics*, Cambridge, 1980.

[32] CHOQUET-BRUHAT, Y. - De WITT-MORETTE, C. - DILLARD-BLEICK, M.: *Analaysis, Manifolds and Physics*, Revised Edition, North-Holland, Amsterdam, 1982.

[33] ARMS, J.M.: Math.Proc.Cambr.Phil.Soc. 90(1981) 361-372: "The structure of the solution set for the Yang-Mills equations".

[34] MARSDEN, J.E.: *Lectures on Geometric Methods in Mathematical Physics*, S.I.A.M., Philadelphia, 1981.

[35] LOOS, H.G.: Jnl.Math.Phys. 8(1967) 2114-2124: "Internal Holonomy Group of Yang-Mills fields".

[36] ZWANZIGER, D.: Nucl.Phys. B209(1982) 336-348: "Nonpertubative modification of the Faddeev-Popov formula and banishment of the naive vacuum".

[37] SINGER, I.M.: Physica Scripta 24(1981) 817-820: "The geometry of orbit space for non-abelian gauge theories".

[38] KOBAYASHI, S. - NOMIZU, K.: *Foundations of Differential Geometry*, Interscience, New York, 1963.

[39] GRIBOV, V.N.: Nucl.Phys. B139(1978) 1-19: "Quantization of non-abelian gauge theories".

[40] BABELON, O. - VIALLET, C.M.: Phys.Lett. 85B(1979) 246-248: "Geometrical interpretation of Faddeev-Popov determinant".

[41] DANIEL, M. - VIALLET, C.M.: Phys.Lett. 76B(1978) 458: "Gauge fixing problem around classical solutions of the Yang-Mills theory".

[42] HANSON, A. - REGGE, T. - TEITELBOIM, C.: *Constrained Hamiltonian Systems*, Academia Nazionale dei Lincei, Rome, 1976.

[43] SYMES, W.W.: Physica 1D(1980) 339-374: "Hamiltonian group actions and integrable systems".

Geometrical Description of the Toda Lattice

A.P.E. ten Kroode
Mathematical Institute, University of Amsterdam

§1. INTRODUCTION

In this seminar we will consider a one-dimensional lattice (chain) of particles, moving under the influence of a nearest neighbour interaction. Such a system may in general be described by a Hamiltonian:

$$H(q,p) = \frac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n-1} \phi(q_{i+1}-q_i) \tag{1}$$

where the p_i's are the momenta of the particles and the q_i's the displacements from their equilibrium positions. ϕ is the interaction potential, which is assumed only to depend on coordinate differences.
The equations of motion corresponding to the Hamiltonian (1) are:

$$\begin{cases} \dot{q}_i = \dfrac{\partial H}{\partial p_i} = p_i \\ \dot{p}_i = -\dfrac{\partial H}{\partial q_i} = \phi'(q_{i+1}-q_i) - \phi'(q_i-q_{i-1}) \end{cases} \tag{2}$$

(where $\phi'(r)$ stands for $\frac{d\phi}{dr}$).

A well-known example is the case of a harmonic interaction. $\phi(r) = \frac{1}{2}Kr^2$.
The system (2) then becomes:

$$\begin{cases} \dot{q}_i = p_i \\ \dot{p}_i = K(q_{i+1}-2q_i+q_{i-1}) \end{cases} \tag{3}$$

(where: $q_o := 0 =: q_{n+1}$, $i = 1,2\ldots n$) or equivalently:

$$\ddot{q}_i = K(q_{i+1}-2q_i+q_{i-1}) \tag{4}$$

This is a set of linear differential equations and the general solution may be written as a superposition of N independent periodic solutions, the so-called *normal oscillation modes*.
We remark at this point that in the limit $n \to \infty$ and the average distance between the particles $\to o$, the discrete label i may be replaced by a con-

tinuous variable x and (4) becomes:

$$\frac{\partial^2 q}{\partial t^2} = K \frac{\partial^2 q}{\partial x^2}; \quad q = q(x,t) \tag{5}$$

(the wave equation).

With the development of the first computers it became possible to study numerically nonlinear interactions; Fermi et al. added for instance cubic and quartic interactions:

$$\phi(r) = \frac{1}{2} Kr^2 + \frac{1}{3} K \alpha r^3 \tag{6a}$$

$$\phi(r) = \frac{1}{2} Kr^2 + \frac{1}{4} K \beta r^4 \tag{6b}$$

To their surprise the computer experiments yielded for suitable initial conditions (not too large energy) again nearly periodic solutions of the equations of motion, belonging to the potentials (6a,b).

These results led Toda[1] to the idea, that there might exist a nonlinear lattice, admitting rigorous periodic solutions. By demanding that the equations (2) have certain periodic solutions he arrived at the potential:

$$\phi(r) = \alpha e^{\beta r} \tag{7}$$

We thus have the Hamiltonian:

$$H(q,p) = \frac{1}{2} \sum_{i=1}^{n} p_i^2 + \alpha \sum_{i=1}^{n-1} e^{\beta(q_{i+1}-q_i)} .$$

If we perform the canonical transformation:

$$\begin{cases} \hat{q}_i = \frac{\beta}{2} q_i + i \,\ell n \frac{4\alpha}{\beta^2} & (\alpha>0) \\ \hat{p}_i = \frac{2}{\beta} p_i \end{cases} \tag{8}$$

we get:

$$H(\hat{q},\hat{p}) = \frac{\beta^2}{4}\left(\frac{1}{2}\sum_{i=1}^{n}\hat{p}_i^2 + \sum_{i=1}^{n-1} e^{2(\hat{q}_{i+1}-\hat{q}_i)}\right)$$

and by rescaling the time $t \longmapsto \frac{\beta^2}{4}t$ in the equations of motion we may as well consider:

$$H(q,p) = \frac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n-1} e^{2(q_{i+1}-q_i)} \tag{9}$$

(9) will be referred to as the Hamiltonian for the *Toda-lattice*.
The associated equations of motion are:

$$\begin{cases} \dot{q}_i = p_i \\ \dot{p}_i = 2\left(e^{2(q_{i+1}-q_i)} - e^{2(q_i-q_{i-1})}\right) \end{cases} \tag{10}$$

$$(i=1,2\ldots n\,; \quad q_o := \infty\,, \quad q_{n+1} := -\infty)$$

or equivalently:

$$\ddot{q}_i = 2\left(e^{2(q_{i+1}-q_i)} - e^{2(q_i-q_{i-1})}\right) \tag{11}$$

Just as in the case of the harmonic interaction we may consider the continuum limit and get:

$$q_t = 6qq_x - q_{xxx}\,; \qquad q = q(x,t) \tag{12}$$

which is the by now famous *Korteweg-de Vries* equation.
The Toda-lattice may thus be considered as a discrete analogon for the KdV-equation and therefore it is not surprising, that we can find *soliton solutions* for (10). The periodic solutions, that led Toda to the exponential lattice in the first place, are in fact the so-called *cnoidal waves*, which are also known from the KdV-equation.
A completely different application of the Toda-equations (10) arises in relation with the study of certain spherical symmetric magnetic monopoles (see [2]).

Flaschka[3] was the first to start an analytical survey of the Toda-equations. He observed that with the definitions:

$$A := \begin{pmatrix} a_1 & b_1 & & \emptyset \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ \emptyset & & b_{n-1} & a_n \end{pmatrix} ; \qquad B := \begin{pmatrix} 0 & b_1 & & \emptyset \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ \emptyset & & b_{n-1} & 0 \end{pmatrix}$$

where $\quad a_i := -p_i \ , \qquad b_i := e^{q_{i+1}-q_i}$ (13)

the equations of motion (10) are equivalent to the matrix differential equation:

$$\dot{A} = [A, B]$$

This equation is known as the *Lax-form* of the equations of motion, after Peter Lax[4] who studied an equation of this type in relation to the Korteweg-de Vries equation. Using the fact that B is antisymmetric, it is not difficult to show, that the eigenvalues of A are constants of the motion. (See e.g. Toda[1]).

Because A is symmetric, it may be diagonalized and thus we get n independent constants of the motion:

$$f_k(q,p) := \frac{1}{k}\mathrm{tr}\, A^k \qquad\qquad k = 1,2\ldots n \tag{15}$$

examples: $f_1(q,p) = \mathrm{tr}\, A = -\sum_{i=1}^{n} p_i =: -P$

$$f_2(q,p) = \frac{1}{2}\mathrm{tr}\, A^2 = \frac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n-1} e^{2(q_{i+1}-q_i)} =: E$$

where P and E are of course the total momentum and total energy of the system. Even more is true; we can prove that the f_k's are in involution:

$$\{f_i, f_j\} = 0 \qquad\qquad \forall i,j \tag{16}$$

(15) and (16) in fact establish, that the Toda-lattice is an example of a so-called completely integrable system. (See e.g. Arnold[5] for a precise definition.)

The intention of this seminar is to place the theory of the Toda-lattice into a geometric framework. Having done this, we will get an interpretation of the Flaschka-transformation (13) and the Lax-equation (14). We will also be able to prove the involution statement (16) and to solve the initial value problem for the Toda-lattice implicitly.

§2. SYMPLECTIC AND POISSON STRUCTURES

In this section we will define two very important notions in a geometric description of dynamical systems namely symplectic structures and Poisson structures. We will see, that these two structures are 'almost' equivalent.

DEFINITION 2.1. A *symplectic structure* is a pair (M,ω) where M is a differentiable manifold and ω a symplectic form, i.e. a closed, nondegenerate 2-form.

DEFINITION 2.2. A *Poisson structure* is a pair $(M,\{\ ,\ \})$ where M is a differentiable manifold and $\{\ ,\ \}$ a Poisson bracket, i.e. a mapping:

$$\{\ ,\ \} : C^\infty(M)\times C^\infty(M) \to C^\infty(M)$$

satisfying the following properties:

1) linearity : $\{f, \lambda g+\mu h\} = \lambda\{f,g\} + \mu\{f,h\}$
2) antisymmetry : $\{f,g\} = -\{g,f\}$
3) Jacobi-identity : $\{f, \{g,h\}\} + \{h, \{f,g\}\} + \{g, \{h,f\}\} = 0$
4) Leibniz-identity: $\{f,gh\} = g\{f,h\} + h\{f,g\}$

REMARK. In Russian usage a Poisson structure is called a *Hamiltonian structure*. In our opinion this terminology is a bit misleading because there is no Hamiltonian present yet.

It is well-known, that if we have a symplectic structure, we can define a Poisson structure by means of the following construction:

- If $f \in C^\infty(M)$, the Hamiltonian vector field X_f is given by the relation:

$$i_{X_f}\omega = df \tag{17}$$

- The Poisson bracket is then given by:

$$\forall f,g \in C^\infty(M),\ x \in M : \{f,g\}(x) := \omega_x(X_f(x),X_g(x)) \tag{18}$$

One now easily verifies the properties 1)-4) of a Poisson bracket. In particular the Jacobi-identity follows from $d\omega = 0$.

We will now consider the reverse procedure; i.e. starting from a Poisson structure we will try to define a symplectic structure.

Given a Poisson bracket we may define $\forall f \in C^\infty(M)$ a mapping:

$$X_f \colon C^\infty(M) \to C^\infty(M)$$

by:

$$X_f(g) = \{g,f\} \qquad \forall g \in C^\infty(M) \tag{19}$$

Because of properties 1) and 4) of a Poisson bracket this is a derivation and therefore it comes from a vector field, which will of course be called the *Hamiltonian vector field associated to* f.

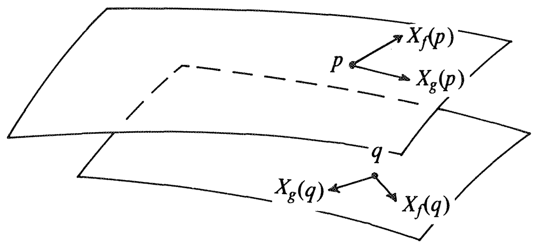
Using the Jacobi-identity, it is easy to show:

$$[X_f, X_g] = X_{\{g,f\}} \tag{20}$$

In words: the collection of Hamiltonian vector fields is involutive.

<u>DEFINITION 2.3</u>. A Poisson structure $(M, \{\ ,\ \})$ is called *regular* if the collection $D_p := \{X_f(p) \mid f \in C^\infty(M)\}$ has constant dimension $\forall p \in M$. It is called *transitive* if: $\dim D_p = \dim M \quad \forall p \in M$.

So if we have a regular Poisson structure, the collection of Hamiltonian fields is integrable according to Frobenius' criterion. M is then foliated by integral manifolds of the Hamiltonian fields and these manifolds all have the same dimension.



EXAMPLE. In $\mathbb{R}^{2n+k}$ with coordinates $(q_1 \ldots q_n, p_1 \ldots p_n, r_1 \ldots r_k)$ we may define a Poisson bracket by:

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

In each point the collection of Hamiltonian fields is spanned by:

$$\frac{\partial}{\partial q_1} \cdots \frac{\partial}{\partial q_n}, \frac{\partial}{\partial p_1} \cdots \frac{\partial}{\partial p_n}.$$

So we have a regular Poisson structure. The integral manifolds are given by:

$$r_1 = c_1, r_2 = c_2 \ldots r_k = c_k.$$

In the case $k = 0$ we have a transitive Poisson structure.

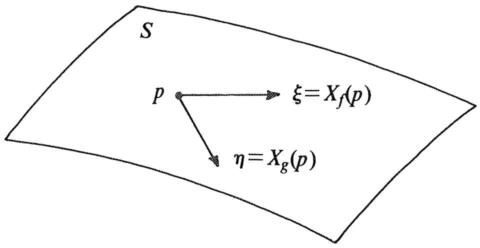
In the general case, however, a Poisson structure need not be regular. We then have the following theorem, which is due to Kirillov[6].

THEOREM 2.1. *If* $(M, \{\ ,\ \})$ *is a Poisson structure, then* M *splits into submanifolds* M_α, *such that each* M_α *is an integral manifold of the Hamiltonian vector fields.*

N.B. We will see an example of a nonregular Poisson structure in §5.

So we arrive at the 'nice' conclusion, that also in the general case M is foliated by integral manifolds of the Hamiltonian fields, but these manifolds may very well have different dimensions.

If S ($= M_\alpha$ for some α) is such an integral manifold, we can equip S with a symplectic form as follows:



$\forall \xi,\eta \in T_pS : \exists f,g \in C^\infty(M)$ such that $\xi = X_f(p)$

$$\eta = X_g(p)$$

define: $$\omega_p(\xi,\eta) := \{f,g\}(p) \tag{21}$$

This is well-defined, because supposing:

$$\xi = X_{f_1}(p) = X_{f_2}(p) \; ; \quad \eta = X_{g_1}(p) = X_{g_2}(p)$$

we have:
$$\begin{aligned}\{f_1,g_1\}(p) &=: X_{g_1}(p)(f_1) \\ &= X_{g_2}(p)(f_1) \\ &= \{f_1,g_2\}(p) \\ &=: -X_{f_1}(p)(g_2) \\ &= -X_{f_2}(p)(g_2) \\ &= \{f_2,g_2\}(p) \quad .\end{aligned}$$

It is not hard to verify, that the 2-form defined above is closed and non-degenerate.

Summarizing we can say, that for any symplectic structure, we can immediately define an associated (regular) Poisson structure. If we start with a Poisson structure $(M, \{\ ,\ \})$, M is foliated in symplectic substructures.

§3. GROUP ACTIONS

Hamiltonian systems often possess symmetries. These symmetries are geometrically best described by the action of a Lie group on the phase space. It is well-known (Noether) that symmetries give rise to conserved quantities. The two most familiar examples are conservation of linear and angular momentum associated to translational and rotational invariance respectively. In general the conserved quantities belonging to certain group actions on the phase space are described by the so-called momentum mapping, which was introduced by Souriau about 1970 and will be defined in this section.

DEFINITION 3.1. Let M be a C^∞-manifold, G a Lie group. A *(left-) action* of G on M is a C^∞-mapping:

$$\phi: G\times M \to M$$

such that:

i) $\forall x \in M : \phi(e,x) = x$

ii) $\forall x \in M,\ g,h \in G : \phi(g,\ \phi(h,x)) = \phi(gh,x)$

REMARK. Instead of $\phi(g,x)$ we often write: $\phi_g(x)$.

If there is on M a symplectic form or a Poisson bracket, we may require that they are preserved by the group action. This is expressed in the following definition:

DEFINITION 3.2. Let (M,ω) be a symplectic structure. A group action is called *symplectic* if:

$$\forall g \in G : \phi_g^*\omega = \omega.$$

Let $(M, \{\ ,\ \})$ be a Poisson structure. A group action is said to be a

Poisson action if:

$$\forall g \in G\,;\ f_1,f_2 \in C^\infty(M) \qquad \phi_g^*\{f_1,f_2\} = \{\phi_g^* f_1, \phi_g^* f_2\}\,.$$

The relation between Poisson actions and symplectic actions is given by the following:

PROPOSITION 3.1.

$$\phi_g^*\omega = \omega \underset{(1)}{\iff} \phi_{g^{-1}*} X_f = X_{\phi_g^* f} \underset{(2)}{\iff} \phi_g^*\{f_1,f_2\} = \{\phi_g^* f_1, \phi_g^* f_2\}\,.$$

PROOF. Abraham & Marsden, chapter 3.

REMARK. The Poisson bracket occurring in this proposition is the Poisson bracket associated to the symplectic form ω. If we had started with a Poisson structure, we still have equivalence (2), but in order for equivalence (1) to hold we must require that the group action leaves invariant the integral manifolds of the Hamiltonian fields.

We will now turn to the infinitesimal description of a group action.

DEFINITION 3.3. Let $\phi\colon G\times M \to M$ be a group action, $\underset{\sim}{g}$ the Lie algebra of G.

For $\xi \in \underset{\sim}{g}$ we define the *Killing-vector field* ξ_M on M or *infinitesimal generator* of the action as:

$$\xi_M(p) := \frac{d}{dt}\,\phi(\exp t\,\xi, p)\Big|_{t=o} \quad (p\in M) \tag{22}$$

N.B. We stress that the subscript M denotes the manifold on which the group acts.

If (M,ω) is a symplectic structure and $\phi\colon G\times M \to M$ a group action, we can ask ourselves if there is a connection between Hamiltonian vector fields and Killing-vector fields.

DEFINITION 3.4. A group action $\phi: G\times M \to M$ is called *Hamiltonian* if there exists a linear mapping:

$$\hat{J}: \underline{g} \to C^{\infty}(M)$$

such that:

$$\xi_M(p) = X_{\hat{J}(\xi)}(p) \quad \forall p \in M \tag{23}$$

We see that in this case any Killing-field ξ_M is generated by the Hamiltonian $\hat{J}(\xi)$. That this situation actually occurs in practice is demonstrated by the following example.

Let ϕ be a group action; $\phi: G\times M \to M$ or equivalently $\phi_g: M \to M \quad \forall g \in G$. We then have the derivative:

$$\phi_{g*}: T_x M \to T_{\phi_g(x)}M$$

and the pullback:

$$\phi_g^*: T^*_{\phi_g(x)}M \to T^*_x M .$$

The composition of two pullbacks is given by:

$$\phi_g^* \circ \phi_h^* = (\phi_h \circ \phi_g)^* = \phi_{hg}^* .$$

The push-forward $\psi_g \equiv \phi^*_{g^{-1}} : T^*_x M \to T^*_{\phi_g(x)}(M)$ obeys:

$$\psi_g \circ \psi_h = \phi^*_{h^{-1}g^{-1}} = \phi^*_{(gh)^{-1}} = \psi_{gh}$$

and therefore $\psi: G\times T^*M \to T^*M$ defined by $\psi_g := \phi^*_{g^{-1}}$ is a (left) action of G on T^*M, called the *induced action* on the cotangent bundle.

We now have the following:

PROPOSITION 3.2. i) ψ *is symplectic with respect to the canonical symplectic structure on* T^*M

ii) ψ *is a Hamiltonian action with Hamiltonian:*

$$\hat{J}(\xi) = \theta(\xi_{T^*M}).$$

PROOF. We first recall for the reader's convenience the definition of the Liouville form θ on $T(T^*M)$.

$$\forall X_\alpha \in T_\alpha(T^*M) \ : \ \theta_\alpha(X_\alpha) := \alpha(\pi_* X_\alpha)$$

where $\alpha \in T^*M$ and $\pi\colon T^*M \to M$ is the canonical projection.

i) By construction of ψ we have commutativity of the following diagram:

$$\begin{array}{ccc} T^*M & \xrightarrow{\ \psi_g\ } & T^*M \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow{\ \phi_g\ } & M \end{array}$$

we then have $\forall \alpha \in T^*M$, $X_\alpha \in T_\alpha(T^*M)$:

$$\begin{aligned} (\psi_g^*\theta)(X_\alpha) &= \theta_{\psi_g(\alpha)}(\psi_{g*}X_\alpha) \\ &:= \psi_g(\alpha)(\pi_* \circ \psi_{g*}X_\alpha) \\ &= \psi_g(\alpha)(\phi_{g*} \circ \pi_* X_\alpha) \\ &= \alpha(\pi_* X_\alpha) \\ &=: \theta_\alpha(X_\alpha) \end{aligned}$$

So $\qquad \psi_g^*\theta = \theta \Rightarrow \psi_g^*\omega = \omega$. ∎

ii) Let $\xi \in \underline{g}$ and $\xi_{T^*M}(\alpha) := \frac{d}{dt}\psi_{\exp t\xi}(\alpha)\Big|_{t=0}$ $\qquad (\alpha \in T^*M)$

we then have: $L_{\xi_{T^*M}}\theta = \frac{d}{dt}(\psi_{\exp t\xi}^*\theta) = 0$.

Using the identity $L_x\theta = i_x(d\theta) + d(i_x\theta)$ we get:

$$0 = L_{\xi_{T^*M}}\theta = i_{\xi_{T^*M}}(d\theta) + d(\theta(\xi_{T^*M}))$$

$$\Rightarrow i_{\xi_{T^*M}}\omega = d(\theta(\xi_{T^*M})) \qquad (\omega := -d\theta)$$

$$\Rightarrow \xi_{T^*M} = X_{\theta(\xi_{T^*M})}.$$

■

For later use we mention the following:

LEMMA 3.3.

$$\pi_*\xi_{T^*M}(\alpha) = \xi_M(\pi(\alpha)) .$$

PROOF. $\pi_*\xi_{T^*M}(\alpha) := \frac{d}{dt}(\pi\circ\psi_{\exp t\,\xi}(\alpha))\Big|_{t=o}$ (see diagram on preceding page)

$$= \frac{d}{dt}\,\phi_{\exp t\,\xi}(\pi(\alpha))\Big|_{t=o}$$

$$=: \xi_M(\pi(\alpha)).$$

■

We are now finally prepared to define the momentum mapping, as was announced in the introduction of this section.

DEFINITION 3.5. Let (M,ω) be a symplectic structure, $\phi: G\times M \to M$ a Hamiltonian action. We then define the *momentum mapping* J *associated to* ϕ by:

$$J: M \to \underline{g}^*$$

$$\forall p \in M,\ x \in \underline{g} : J(p)(x) := \hat{J}(x)(p) \tag{25}$$

N.B. $J(p) \in \underline{g}^*$ because $\hat{J}: \underline{g} \to C^\infty(M)$ is linear.

That the momentum mapping J of a Hamiltonian group action is indeed a very useful geometric tool to describe constants of the motion, will be clear from the following:

PROPOSITION 3.4. *Let* $\phi: G\times M \to M$ *be an action and* $\psi: g\times T^*M \to T^*M$ *the induced action. If* $H: T^*M \to \mathbb{R}$ *is a Hamiltonian invariant under* ψ, i.e. $H\circ\psi = H$, *the momentum mapping* $J: T^*M \to \underline{g}^*$ *associated to* ψ *is constant along the integral curves of* X_H.

PROOF. Define $\gamma: t \mapsto \gamma(t) \in T^*M$ by: $\dot\gamma(t) = X_H(\gamma(t))$

$$\begin{aligned}\forall\xi\in\underline{g} : \frac{d}{dt} J(\gamma(t))(\xi) &:= \frac{d}{dt}\hat{J}(\xi)(\gamma(t)) \\ &= X_H(\gamma(t))[\hat{J}(\xi)] \\ &= \{\hat{J}(\xi), H\}(\gamma(t)) \\ &= -X_{\hat{J}(\xi)}(\gamma(t))[H] \\ &= -\xi_{T^*M}(\gamma(t))[H] \\ &= -\frac{d}{ds} H\circ\psi_{\exp s\,\xi}(\gamma(t))\Big|_{s=o} \\ &= 0.\end{aligned}$$

■

EXAMPLES.

1) $M = \mathbb{R}^3$ coordinates $\vec{q} = (q_1,q_2,q_3)$

$T^*M = T^*\mathbb{R}^3$ coordinates $(\vec{q},\vec{p})$.

Choose $G = \mathbb{R}^3$, the 3-dimensional Euclidean group:

$$\phi_{\vec{x}}(\vec{q}) := (\vec{q}+\vec{x})$$

$$\forall\vec{\xi} = (\xi_1,\xi_2,\xi_3)\in\underline{g} = \mathbb{R}^3 .$$

$$\begin{aligned}J(\vec{q},\vec{p})(\vec{\xi}) &:= \hat{J}(\vec{\xi})(\vec{q},\vec{p}) \\ &= \theta_{(\vec{q},\vec{p})}(\vec{\xi}_{T^*\mathbb{R}^3}(\vec{q},\vec{p})) && \text{(Prop.3.2 ii))} \\ &:= (\vec{q},\vec{p})(\pi_*\vec{\xi}_{T^*\mathbb{R}^3}(\vec{q},\vec{p})) \\ &= (\vec{q},\vec{p})(\vec{\xi}_{\mathbb{R}^3}(\vec{q})) && \text{(Lemma 3.3)} \\ &= \vec{p}\cdot\vec{\xi} .\end{aligned}$$

So $J(\vec{q},\vec{p}) = \vec{p}$ and for translation invariant Hamiltonians the linear momentum will be conserved.

2) $M = \mathbb{R}^3$, $G = SO(3)$

$\forall A \in SO(3) : \phi_A(\vec{q}) := A\vec{q}$

$\underset{=}{g} = \{X \in M(3,\mathbb{R}) \mid X^T = -X\}$

basis of $\underset{=}{g}$: $E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$

$$E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

$$E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\forall \xi \in \underset{=}{g} \; : \; \xi = \xi_1 E_1 + \xi_2 E_2 + \xi_3 E_3 = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix}$$

$$\xi_{\mathbb{R}^3}(\vec{q}) := \frac{d}{dt}\, \phi_{\exp t\xi}(\vec{q})\Big|_{t=o} = \frac{d}{dt} \exp t\xi\, \vec{q}\Big|_{t=o}$$

$$= \xi\vec{q} \,.$$

As in example (1) we find:

$$J(\vec{q},\vec{p})(\xi) = p \cdot (\xi\vec{q}) = \vec{\xi} \cdot (\vec{q} \wedge \vec{p})$$

where $\vec{\xi} := (\xi_1, \xi_2, \xi_3)$.

So: $J(\vec{q},\vec{p}) = \vec{q} \wedge \vec{p}$ and for rotation invariant Hamiltonians the angular momentum will be conserved.

3) We now come to the most important example for our purposes.

Choose $M = G$, a Lie group. G can act on itself by means of left and right multiplications.

$$L_g(h) = gh \,, \quad R_g(h) = hg.$$

The hereby induced actions on T^*G will be denoted by:

$$\bar{L}_g = L^*_{g^{-1}}, \quad \bar{R}_g = R^*_{g^{-1}}.$$

The momentum mapping for the left-action becomes: $\alpha \in T^*G$, $\xi \in \underline{g}$:

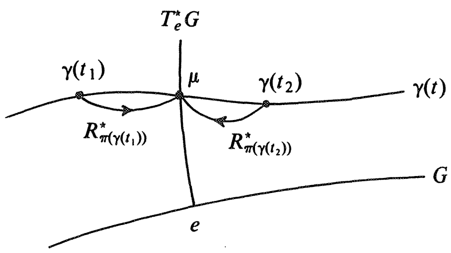
$$\begin{aligned} J^L(\alpha)(\xi) &:= \hat{J}^L(\xi)(\alpha) \\ &= \theta_\alpha(\xi^L_{T^*G}(\alpha)) \\ &= \alpha(\pi_*\xi^L_{T^*G}(\alpha)) \\ &= \alpha(\xi^L_G(\pi(\alpha))) \\ &= \alpha(\frac{d}{dt} L_{\exp t\,\xi}(\pi(\alpha)))\Big|_{t=o} \\ &= \alpha(R_{\pi(\alpha)*}\xi) \\ &= (R^*_{\pi(\alpha)}\alpha)(\xi). \end{aligned}$$

So: $$J^L(\alpha) = R^*_{\pi(\alpha)}\alpha \tag{26a}$$

For the right-action we will of course find:

$$J^R(\alpha) = L^*_{\pi(\alpha)}\alpha \tag{26b}$$

We see that in both cases the form $\alpha \in T^*_{\pi(\alpha)}G$ is pulled back to the fibre over the identity $T^*_eG \simeq \underline{g}^*$



The mathematical interpretation of the conserved quantity is shown in the picture above; $\gamma(t)$ is the integral curve of a left-invariant Hamiltonian.

If we pullback an arbitrary point on $\gamma(t)$ to T_e^*G using the induced right-action, we will always end up in the same point μ.

As a physical application of this construction we mention the case of a rigid body, where the configuration space is indeed a Lie group, namely $G = SO(3)$. The conservation laws ($26^{a,b}$) turn out to be the so-called *Euler conservation laws*. (Abraham & Marsden, Ch.4)

§4. SYMMETRY AND REDUCTION

When there are symmetries present in a Hamiltonian system, the phase space can be reduced; that is: a number of variables can be eliminated. The result is called the reduced phase space, which contains the essential dynamics. The procedure by which it was obtained is called reduction. In this section we will describe the reduction procedure, thereby using the notions of a Hamiltonian group action and its associated momentum mapping, as defined in the previous section.

The setting is as follows:
Let M be a C^∞-manifold, K a Lie group, $\phi\colon K\times M \to M$ a group action.
We may divide out M with respect to the K-action and suppose:

- $L \equiv M/K$ is a C^∞-manifold
- the projection $f\colon M \to L$ is a C^∞-surjection
- $f_*\colon TM \to TL$ is a C^∞-surjection

In the previous section we saw, that the induced action on the cotangent bundle:

$$\psi\colon K\times T^*M \to T^*M$$

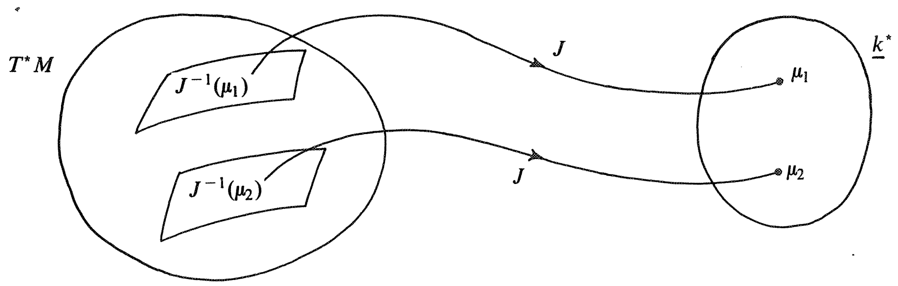
is Hamiltonian and therefore possesses a momentum mapping:

$$J\colon T^*M \to \underline{k}^* .$$

We also saw, that this momentum mapping is conserved along the integral curves of a Hamiltonian vector field associated to a K-invariant Hamiltonian. Therefore it seems rather natural to study surfaces in T^*M on which J has a constant value $\mu \in \underline{k}^*$:

$$J^{-1}(\mu) := \{\alpha \in T^*M \mid J(\alpha) = \mu\} .$$

We will assume that $J^{-1}(\mu)$ is a submanifold of T^*M.



The following proposition is now almost trivial:

PROPOSITION 4.1. *If* $g \in C^\infty(T^*M)$ *is* K-*invariant and* F_t *is the flow associated to the Hamiltonian vector field* X_g, *we have:*

$$F_t(J^{-1}(\mu)) \subset J^{-1}(\mu) .$$

PROOF. $\forall\alpha \in J^{-1}(\mu) \ : \ J(F_t(\alpha)) = J(\alpha) = \mu \ \Rightarrow \ F_t(\alpha) \in J^{-1}(\mu)$. ∎

We see, that if we have a Hamiltonian system with symmetry, or more precisely: if we have a K-invariant Hamiltonian on T^*M, the motion is restricted to surfaces $J^{-1}(\mu)$. For our purposes it is sufficient to consider $J^{-1}(0)$. (See: Abraham & Marsden, Ch.4 for the general case.) We then have:

LEMMA 4.2.

$$\alpha \in J^{-1}(0) \iff \alpha(\xi_M(\pi(\alpha))) = 0 \quad \forall\xi \in \underline{k} .$$

PROOF. $J(\alpha)(\xi) := \hat{J}(\xi)(\alpha)$

$$= \theta_\alpha(\xi_{T^*M}(\alpha)) \qquad \text{(prop. 3.2ii)}$$

$$:= \alpha(\pi_*\xi_{T^*M}(\alpha))$$

$$= \alpha(\xi_M(\pi(\alpha))). \qquad \text{(lemma 3.3)}$$

∎

If $f\colon M \to M/K = L$ is again the projection as defined above, it is easy to see that:

$$\ker f_*\big|_x = \{\xi_M(x) \mid \xi \in \underline{k}\}$$

(So: $\alpha \in J^{-1}(0) \Longleftrightarrow \alpha(\ker f_*\big|_{\pi(\alpha)}) = \{0\})$.

Now f is of course K-invariant:

$$f \circ \phi_k = f \qquad \forall k \in K$$

differentiating both sides in $x \in M$ we get:

$$f_*\big|_{\phi_k(x)} \circ \phi_{k^*}\big|_x = f_*\big|_x$$

$$\Longrightarrow \quad x \in \ker f_*\big|_x \Longleftrightarrow \phi_{k^*}x \in \ker f_*\big|_{\phi_k(x)}$$

or equivalently: $\phi_{k^*}\ker f_*\big|_x = \ker f_*\big|_{\phi_k(x)}$.

It is now easy to prove the following:

PROPOSITION 4.3. $J^{-1}(0)$ *is invariant under the* K-*action*.

PROOF. According to lemma 4.2 we have to prove that:

$$(\psi_k\alpha)(\xi_M(\pi(\psi_K\alpha))) = 0 \qquad \forall \alpha \in J^{-1}(0), \quad k \in K, \quad \xi \in \underline{k}.$$

$$(\psi_K\alpha)(\xi_M(\pi(\psi_k\alpha))) = \alpha(\phi_{k^{-1}*}\xi_M(\pi(\psi_k\alpha)))$$
$$= \alpha(\phi_{k^{-1}*}\xi_M(\phi_k\pi(\alpha)))$$

$$\xi_M(\phi_k\pi(\alpha)) \in \ker f_*\big|_{\phi_k\pi(\alpha)} \longrightarrow \phi_{k^{-1}*}\xi_M(\phi_k\pi(\alpha)) \in \ker f_*\big|_{\pi(\alpha)}$$
$$\longrightarrow \alpha(\phi_{k^{-1}*}\xi_M(\phi_k\pi(\alpha))) = 0. \quad \blacksquare$$

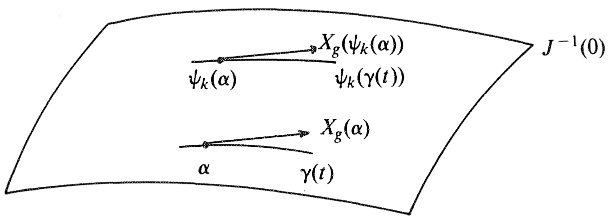
If we have in $J^{-1}(0)$ a solution curve $\gamma(t)$ of a K-invariant Hamiltonian, we can ask ourselves how K acts on $\gamma(t)$.

PROPOSITION 4.4. K *transforms a solution curve of a* K-*invariant Hamiltonian* $g \in C^\infty(T^*M)$ *into another solution curve of* g.

PROOF. The action ψ: $K \times T^*M \to T^*M$ is induced and therefore symplectic (prop. 3.2). According to proposition 3.1 we then have:

$$\psi_{k*}X_g(\alpha) = X_{\psi^*_{k^{-1}}g}(\psi_k(\alpha))$$
$$= X_g(\psi_k(\alpha)) \qquad \text{(g is K-invariant)}$$

From this the proposition is immediate. ■



So once we have a solution curve $\gamma(t)$ in $J^{-1}(0)$, we can generate others by performing K-transformations on $\gamma(t)$.

REMARK. Sometimes there is no physical difference between $\gamma(t)$ and $\psi_k\gamma(t)$. The group K is then called a *gauge group* and the solutions γ and $\psi_k\gamma$ are called *gauge equivalent*.

We are now going to construct a mapping $\chi\colon J^{-1}(0) \to T^*L$, projecting integral curves of K-invariant Hamiltonians in $J^{-1}(0)$ on integral curves of certain associated Hamiltonians in T^*L. It will appear, that $\gamma(t)$ and $\psi_k\gamma(t)$ have the same image under χ. Therefore T^*L contains the essential dynamics and is often called the *reduced phase space*.

Construction.

$$\chi\colon \alpha \in J^{-1}(0) \longmapsto \hat{\alpha} \in T^*_{f(\pi(\alpha))}L$$

$$\forall \hat{\xi} \in T_{f(\pi(\alpha))}L \quad : \quad \exists \xi \in T_{\pi(\alpha)}M \quad \text{such that} \quad \hat{\xi} = f_*\xi \,.$$

Now define: $\hat{\alpha}(\hat{\xi}) := \alpha(\xi)$ (27)

This is well-defined because supposing:

$$\hat{\xi} = f_*\xi = f_*\eta \qquad\qquad \xi,\eta \in T_{\pi(\alpha)}M$$

we have:

$$\begin{aligned} f_*(\xi-\eta) = 0 &\iff \xi - \eta \in \ker f_* \\ &\iff \alpha(\xi-\eta) = 0 \qquad\qquad \text{(lemma 4.2)} \\ &\iff \alpha(\xi) = \alpha(\eta)\,. \end{aligned}$$

The mapping χ is surjective because:

$$\forall \hat{\alpha} \in T^*_{f(x)}L \quad : \quad \hat{\alpha}\circ f_*\big|_x \in T^*_xM \quad \text{is a preimage of } \hat{\alpha} \text{ under } \chi\,.$$

(see (27))

χ is however not injective because if $f(x) = f(y)$, $\hat{\alpha} \circ f_*|_y$ is also a preimage of $\hat{\alpha}$.

We will now show that there is a simple relation between any two preimages.

$$f(x) = f(y) \Rightarrow \exists k \in K \quad \text{such that} \quad y = \phi_k(x)$$

because of the invariance of f under K

$$f \circ \phi_k = f$$

we have again:

$$f_*|_{y=\phi_k(x)} \circ \phi_{k*}|_x = f_*|_x .$$

So: $$\hat{\alpha}\circ f_*|_x = \hat{\alpha}\circ f_*|_y \circ \phi_{k*}|_x = \phi_k^*(\hat{\alpha}\circ f_*|_y) .$$

So we have: $\forall \alpha,\beta \in J^{-1}(0) : \hat{\alpha} = \hat{\beta} \Rightarrow \exists k \in K$ such that $\alpha = \phi_k^* \beta$.

A function $g \in C^\infty(T^*M)$, invariant under the K-action, induces in a very natural way a function $\hat{g} \in C^\infty(T^*L)$ by means of:

$$\hat{g}(\hat{\alpha}) := g(\alpha) \tag{28}$$

<u>REMARK</u>. If g is the Hamiltonian on T^*M $\hat{g}$ is called the *reduced Hamiltonian* on T^*L.

Conversely, any function $\hat{g} \in C^\infty(T^*L)$ comes from a K-invariant function $g \in C^\infty(T^*M)$.

We now have the following

<u>PROPOSITION 4.5</u>. i) *If* θ, ψ *are the canonical* 1-*forms on* T^*M *and* T^*L *respectively, we have:*

$$\chi^*\psi = \theta|_{TJ^{-1}(0)}$$

ii) If $g \in C^{\infty}(T^{*}M)$, $\hat{g} \in C^{\infty}(T^{*}L)$ are as above, we have:

$$\chi_{*}X_{g}(\alpha) = X_{\hat{g}}(\hat{\alpha}) \tag{29}$$

iii) $$\{\hat{g},\hat{h}\}(\hat{\alpha}) = \{g,h\}(\alpha) \tag{30}$$

PROOF. i) The following diagram commutes:

$$\begin{array}{ccc} J^{-1}(0) \subset T^{*}M & \xrightarrow{\ \chi\ } & T^{*}L \\ \downarrow{\scriptstyle \pi^{M}} & & \downarrow{\scriptstyle \pi^{L}} \\ M & \xrightarrow{\ f\ } & L \end{array}$$

because: $\alpha \in J^{-1}(0) \overset{\chi}{\longmapsto} \hat{\alpha} \in T^{*}_{f(\pi^{M}(\alpha))}L \overset{\pi^{L}}{\longmapsto} f(\pi^{M}(\alpha)) \in L$

$$\alpha \in J^{-1}(0) \xrightarrow{\pi^{M}} \pi^{M}(\alpha) \in M \overset{f}{\longmapsto} f(\pi^{M}(\alpha)) \in L \ .$$

We now have $\forall X \in T_{\alpha}J^{-1}(0)$.

$$\begin{aligned} (\chi^{*}\psi)_{\alpha}(X) &= \psi_{\hat{\alpha}}(\chi_{*}X) \\ &:= \hat{\alpha}(\pi^{L}_{*}\circ\chi_{*}X) \\ &= \hat{\alpha}(f_{*}\circ\pi^{M}_{*}X) \\ &:= \alpha(\pi^{M}_{*}X) \\ &=: \theta_{\alpha}(X) \ . \end{aligned}$$

So: $$\chi^{*}\psi = \theta\Big|_{TJ^{-1}(0)} \ . \qquad \blacksquare$$

N.B. If $\rho := -d\psi$ and $\omega := -d\theta$ are the associated symplectic forms, we also have of course:

$$\chi^{*}\rho = \omega\Big|_{TJ^{-1}(0)\times TJ^{-1}(0)} \ .$$

ii) Note that:

$$\hat{g}(\hat{\alpha}) = \hat{g} \circ \chi(\alpha) = g(\alpha)$$

so: $g = \hat{g} \circ \chi = \chi^* \hat{g}$

$$\begin{aligned}
\chi^*(i_{\chi_* X_g(\alpha)} \rho) &= i_{X_g(\alpha)} \chi^* \rho \\
&= i_{X_g(\alpha)} \omega \\
&= dg\big|_{\alpha} \\
&= \chi^*(d\hat{g}\big|_{\hat{\alpha}}) \\
&= \chi^*(i_{X_{\hat{g}}(\hat{\alpha})} \rho) \\
\Longrightarrow \quad i_{\chi_* X_g(\alpha)} \rho &= i_{X_{\hat{g}}(\hat{\alpha})} \rho \\
\Longrightarrow \quad \chi_* X_g(\alpha) &= X_{\hat{g}}(\hat{\alpha}) . \qquad \blacksquare
\end{aligned}$$

iii)

$$\begin{aligned}
\{\hat{g},\hat{h}\}(\hat{\alpha}) &= X_{\hat{h}}(\hat{\alpha})[\hat{g}] \\
&= \chi_* X_h(\alpha)[\hat{g}] \\
&= X_h(\alpha)[\hat{g} \circ \chi] \\
&= X_h(\alpha)[g] \\
&= \{g,h\}(\alpha) . \qquad \blacksquare
\end{aligned}$$

Point ii) of this proposition tells us that integral curves of a K-invariant Hamiltonian $g \in C^\infty(T^*M)$ are projected on integral curves of the reduced Hamiltonian $\hat{g} \in C^\infty(T^*L)$.
Furthermore $\gamma(t)$ and $\psi_k \gamma(t)$ have the same image under χ.

Finally we discuss a very special case of the above construction.
Let $M = G$ be a Lie group and assume that G may be decomposed in two Lie

subgroups K and L i.e.:

$\forall Q \in G : \exists! Z \in K, \exists! X \in L$ such that:

$Q = ZX$.

We simply write: $G = KL$.

We then have

$M/K = G/K = L$ and the notation used above may be maintained.

REMARK. We saw that functions on $T^*M \equiv T^*G$, invariant under the (left) action of K on T^*G, induce functions on T^*L.
Often we simply start with functions on T^*G, invariant under the left-action of the whole group G on T^*G and call them shortly *left-invariant*.

We conclude this section with a brief summary:

- A symmetry is described by the (induced) action of a Lie group K on the phase space T^*M
- The Hamiltonian is K-invariant and the solution curves lie in the surfaces $J^{-1}(\mu)$
- $J^{-1}(0)$ is invariant under K and solutions in $J^{-1}(0)$ may be transformed into each other by K-actions
- $\chi: J^{-1}(0) \to T^*L$ projects solutions of the Hamiltonian g on solutions of the reduced Hamiltonian $\hat{g}$; $\gamma(t)$ and $\psi_k\gamma(t)$ have the same image under χ.

§5. KIRILLOV'S POISSON STRUCTURE

The intention of this section is to introduce a Poisson structure on the dual $\underline{g}^*$ of any (finite-dimensional) Lie algebra $\underline{g}$ of a Lie group G. According to Weinstein[7] this structure has already been found by Lie himself about 1890. It was rediscovered in the sixties by Kirillov (Study of unitary representations of nilpotent Lie groups), Kostant and Souriau (geometric quantization.)

The set-up will be rather similar to the one in the preceding paragraph. Again we are going to divide out the symmetries of a Hamiltonian system. However, in this section we will take for the phase space of our Hamiltonian system an arbitrary Poisson manifold. (So in general not a cotangent bundle.) The symmetry is again described by a group action but in general this action need not be Hamiltonian.

To be specific:
Let $(M, \{\ ,\ \})$ be a Poisson manifold and $\phi: G\times M \to M$ a Poisson action of G on M.

We will first formulate an analogue of proposition 4.4:

<u>PROPOSITION 5.1</u>. *Let* $f \in C^\infty(M)$ *be a* G*-invariant Hamiltonian. Then* G *transforms a solution curve of* f *into another solution curve of* f *.*

<u>PROOF</u>. The G-action is assumed to be a Poisson action and therefore we have according to prop.3.1:

$$\begin{aligned} \phi_{g*}X_f &= X_{\phi^*_{g^{-1}}f} \\ &= X_f \qquad\qquad (f \text{ is } G\text{-invariant}) \end{aligned}$$

From this the proposition is immediate. ∎

We are now going to divide out the G-symmetry and we will assume:

- M/G is a C^∞-manifold
- $P: M \to M/G$ is a submersion.

N.B. Note that we are dividing out the whole phase space with respect to the G-action instead of only a part $(J^{-1}(0))$ of it, as was the case in §4.

If $f,g \in C^\infty(M)$ are invariant under the G-action:

$$f \circ \phi_k = f, \quad g \circ \phi_k = g \quad \forall k \in G$$

they will induce functions $\tilde{f}, \tilde{g} \in C^\infty(M/G)$ by:

$$\tilde{f}(P(p)) := f(p) , \quad \tilde{g}(P(p)) := g(p) .$$

Because the G-action is assumed to be a Poisson action, the Poisson bracket between two invariants will be invariant itself:

$$\begin{aligned} \{f,g\} \circ \phi_k &= \phi_k^*\{f,g\} \\ &= \{\phi_k^* f, \phi_k^* g\} \\ &= \{f,g\} . \end{aligned}$$

This enables us to define a Poisson bracket on M/G by:

$$\{\tilde{f}, \tilde{g}\}(P(p)) := \{f,g\}(p) \tag{31}$$

From this definition we see:

$$\begin{aligned} X_{\tilde{g}}(P(p))[\tilde{f}] &:= \{f,g\}(p) \\ &:= (p) \\ &= X_g(p)[f] \\ &= X_g(p)[\tilde{f} \circ P] \\ &= (P_* X_g(p))[\tilde{f}] \qquad \forall \tilde{f} \in C^\infty(M/G) \end{aligned}$$

$$\Longrightarrow \quad P_* X_g(p) = X_{\tilde{g}}(P(p)) \tag{32}$$

In words this means, that the solution curves of a G-invariant Hamiltonian g on M are projected on the solution curves of the induced Hamiltonian $\tilde{g}$ on M/G by P. It is obvious that $\gamma(t)$ and $\phi_g \gamma(t)$ have the same image under P.

We will now construct an example of this rather general situation by choosing:

$$M = T^*G , \qquad \phi_g := \bar{L}_g \equiv L^*_{g^{-1}} .$$

The action of G on T^*G is induced and therefore a Poisson action. (see prop. 3.2)

In order to consider the space T^*G/G we take an $\alpha \in T^*G$. The orbit O_α under the action $\bar{L}_g$ becomes:

$$O_\alpha := \{\beta \in T^*G \mid \beta = L^*_{g^{-1}}\alpha \quad \text{for some} \quad g \in G\}$$

and this corresponds precisely to a left-invariant one-form!
So we may conclude: $T^*G/G \equiv \underline{g}^*$. It is obvious, that this is a manifold. Under the identification $\underline{g}^* \simeq T^*_eG$ the projection

$$P: T^*G \to \underline{g}^*$$

simply becomes the pullback to the fibre over the identity by means of left-translation

$$P(\alpha) = L^*_{\pi(\alpha)}\alpha \tag{33}$$

Using the above construction, we can now define a Poisson bracket on $\underline{g}^*$:

$$\forall \tilde{f}, \tilde{g} \in C^\infty(\underline{g}^*) , \quad \mu \in \underline{g}^* \simeq T^*_eG :$$

$$\{\tilde{f}, \tilde{g}\}(\mu) := \{f,g\}(\mu) \tag{34}$$

where $f = \tilde{f} \circ P$, $g = \tilde{g} \circ P$ are the corresponding left-invariant functions on T^*G.

This bracket is called *Kirillov's Poisson bracket.*

In order to get an explicit expression for this bracket, we will need the following defintions.

DEFINITION 5.1. Let M be a C^∞-manifold, $f \in C^\infty(M)$, $p \in M$. The *gradient of* f *in* p is defined as the second component of $df|_p$,

i.e. $\quad df|_p = (p, \nabla f(p))$.

If $(x_1,\ldots,x_n)$ are local coordinates around p we simply get:

$$\nabla f(p) = \left(\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_n}(p) \right).$$

If $M \equiv \underline{g}^*$, $\tilde{f} \in C^\infty(\underline{g}^*)$, $\alpha \in \underline{g}^*$ we have:

$$d\tilde{f}|_\alpha = (\alpha, \nabla \tilde{f}(\alpha))$$

where $\nabla \tilde{f}(\alpha)$ is considered to be an element of $T^*_\alpha g^* \simeq \underline{g}^{**} \simeq \underline{g}$.

So: $\quad \nabla \tilde{f}\colon \alpha \in g^* \longmapsto \nabla \tilde{f}(\alpha) \in \underline{g}$.

DEFINITION 5.2. Let G be a Lie group, $\underline{g}$ its algebra. The *conjugation action* of G on itself is defined by:

$$I_g\colon G \to G$$

$$\forall g,h \in G\colon I_g(h) := ghg^{-1} \tag{35}$$

In other words: $I_g := L_g \circ R_{g^{-1}}$.

The derivative of this mapping at e is an action of G on $\underline{g}$, called the *adjoint action*:

$$Ad_g : \underline{g} \simeq T_e G \to \underline{g}$$

$$Ad_g := I_{g*}\big|_e = (L_{g*} \circ R_{g^{-1}*})\big|_e \tag{36}$$

Associated to this action we define the *coadjoint action* of G on $\underline{g}^*$ by

$$Ad'_g : \underline{g}^* \to \underline{g}^*$$

$$Ad'_g := (Ad_{g^{-1}})^* \equiv L^*_{g^{-1}} \circ R^*_g \tag{37}$$

We can now compute the associated Killing vector fields: (See e.g. Abraham & Marsden, Ch. 4.1)

$$\forall \xi, \eta \in \underline{g}, \quad \alpha \in \underline{g}^* :$$

$$\xi_{\underline{g}}(\eta) := \frac{d}{dt} Ad_{\exp t\,\xi}(\eta)\Big|_{t=o} = [\xi, \eta] \tag{38}$$

and:

$$\xi_{\underline{g}^*}(\alpha) := \frac{d}{dt} Ad'_{\exp t\,\xi}(\alpha)\Big|_{t=o}$$

and indicate them by:

$$\xi_{\underline{g}} \equiv ad_\xi : \underline{g} \to \underline{g}$$

$$ad_\xi(\eta) = [\xi, \eta]$$

$$\xi_{\underline{g}^*} \equiv ad'_\xi : \underline{g}^* \to \underline{g}^* .$$

We have:

$$\forall \alpha \in \underline{g}^*, \quad \xi, \eta \in \underline{g} : \qquad ad'_\xi(\alpha)(\eta) := \frac{d}{dt} Ad'_{\exp t\,\xi}(\alpha)\Big|_{t=o}(\eta)$$

$$= \alpha\left(\frac{d}{dt} Ad_{\exp - t\xi}(\eta)\Big|_{t=o}\right)$$

$$= -\alpha(ad_\xi(\eta))$$

$$= [(-ad_\xi)^* \alpha](\eta) .$$

So: $$ad'_\xi = -(ad_\xi)^* \tag{39}$$

We will now compute explicitly Kirillov's Poisson bracket on $\underline{g}^*$. The calculations are rather technical and therefore the reader may wish to skip them and just look at the result (42).

$\forall \alpha \in T^*_{\pi(\alpha)}G$, f left-invariant:

$$\begin{aligned} f(\alpha) &= \tilde{f}(P(\alpha)) \\ &= \tilde{f}(L^*_{\pi(\alpha)}\alpha) \\ &= \tilde{f}(J^R(\alpha)) \\ &= (J^{R*}\tilde{f})(\alpha) \end{aligned}$$

where J^R is the momentum mapping associated to the right-action of G on T^*G as discussed in §3 example 3.

We will now first compute J^* for an arbitrary momentum mapping.

$$J: M \to \underline{g}^*$$

$$J_*: T_pM \to T_{J(p)}\underline{g}^* \simeq \underline{g}^*$$

$$J^*: T^*_{J(p)}\underline{g}^* \simeq \underline{g}^{**} \simeq \underline{g} \to T^*_pM .$$

So if $\xi \in \underline{g} \simeq T^*_{J(p)}\underline{g}^*$, then $J^*\xi \in T^*_pM$.

$\forall v \in T_pM$ with: $v = \dot{\gamma}(0)$, $\gamma(0) = p$:

$$\begin{aligned} J^*\xi(v) &= \xi(J_*v) \\ &= \frac{d}{dt}\, \xi(J(\gamma(t)))\Big|_{t=o} \\ &= \frac{d}{dt}\, \hat{J}(\xi)(\gamma(t))\Big|_{t=o} \qquad \text{(form. (25))} \\ &= d(\hat{J}(\xi))\Big|_p(v) . \end{aligned}$$

So we may conclude:

$$J^{*}\xi = d(\hat{J}(\xi)) \quad (\xi \in \underline{g} \simeq T^{*}_{J(p)}\underline{g}^{*}) \tag{40}$$

We are now prepared for the following:

<u>PROPOSITION 5.2.</u> *let* $f \in C^{\infty}(T^{*}G)$ *be left-invariant,* $\tilde{f}$ *the corresponding function on* $\underline{g}^{*}$. *Then* $\forall\mu \in T^{*}_{e}G \simeq \underline{g}^{*}$ *the Hamiltonian vector field associated to* f *at* μ *is given by:*

$$\begin{aligned} X_f(\mu) &= \frac{d}{dt}(\bar{R}_{\exp t\,\nabla\tilde{f}(\mu)}(\mu))\Big|_{t=o} \\ &= [\nabla\tilde{f}(\mu)]^{R}_{T^{*}G}\Big|_{\mu} \end{aligned} \tag{41}$$

In words: At the fibre over the identity $T^{*}_{e}G$ the Hamiltonian vector field X_f, associated to a left-invariant Hamiltonian f, coincides with the Killing vector field $[\nabla\tilde{f}(\mu)]^{R}_{T^{*}G}$ of the right-action of G on $T^{*}G$.

<u>PROOF.</u>

$$\begin{aligned} i_{X_f(\mu)}\omega_\mu &= df\Big|_{\mu} \\ &= d(J^{R^{*}}\tilde{f})\Big|_{\mu} \\ &= J^{R^{*}}(d\tilde{f}\Big|_{\mu}) \\ &= J^{R^{*}}((\mu, \nabla\tilde{f}(\mu))) \\ &\qquad (\nabla\tilde{f}(\mu) \in g \simeq T^{*}_{J^{R}(\mu)}\underline{g}^{*} \\ &= d(\hat{J}^{R}(\nabla\tilde{f}(\mu)))\Big|_{\mu} \qquad \text{(form. (40))}. \\ &= i_{X_{\hat{J}^{R}(\nabla\tilde{f}(\mu))}\big|_{\mu}}\omega_\mu \end{aligned}$$

$$\Rightarrow \quad X_f(\mu) = X_{\hat{J}^{R}(\nabla\tilde{f}(\mu))}\Big|_{\mu} = [\nabla\tilde{f}(\mu)]^{R}_{T^{*}G}\Big|_{\mu}. \quad \blacksquare$$

It is now easy to compute Kirillov's Poisson bracket:

$$\{\tilde{f},\tilde{g}\}(\mu) := \{f,g\}(\mu)$$

$$= X_g(\mu)[f]$$

$$= \frac{d}{dt} f(\bar{R}_{\exp t\nabla\tilde{g}(\mu)}(\mu))\Big|_{t=o}$$

$$= \frac{d}{dt} \tilde{f}(\bar{L}_{\exp -t\nabla\tilde{g}(\mu)} \circ \bar{R}_{\exp t\nabla\tilde{g}(\mu)}(\mu))\Big|_{t=o}$$

$$= \frac{d}{dt} \tilde{f}(Ad'_{\exp -t\nabla\tilde{g}(\mu)}(\mu))\Big|_{t=o}$$

$$= -[ad'_{\nabla\tilde{g}(\mu)}(\mu)]\,(\nabla\tilde{f}(\mu))$$

$$= -\mu([\nabla\tilde{f}(\mu), \nabla\tilde{g}(\mu)]).$$

Result: Kirillov's Poisson bracket is given by:

$$\{\tilde{f},\tilde{g}\}(\mu) = -\mu([\nabla\tilde{f}(\mu), \nabla\tilde{g}(\mu)]) \tag{42}$$

As was explained in §2, we can, given a Poisson structure, always define a symplectic structure on the integral manifolds of the Hamiltonian vector fields. Therefore we will now study those integral manifolds.

The Hamiltonian vector fields on $\underline{g}^*$ are given by:

$$\forall \tilde{f},\tilde{g} \in C^\infty(\underline{g}^*),\ \alpha \in \underline{g}^*: \quad X_{\tilde{f}}(\alpha)[\tilde{g}] = -\{\tilde{f},\tilde{g}\}(\alpha)$$

$$= \alpha([\nabla\tilde{f}(\alpha), \nabla\tilde{g}(\alpha)])$$

$$= -(ad'_{\nabla\tilde{f}(\alpha)}\alpha)(\nabla\tilde{g}(\alpha)) .$$

In coordinates this reads:

$$\sum_i X^i_{\tilde{f}}(\alpha)\, \nabla\tilde{g}(\alpha)^i = \sum_i (-ad'_{\nabla\tilde{f}(\alpha)}\alpha)^i\, \nabla\tilde{g}(\alpha)^i .$$

If $\alpha(t) \subset \underline{g}^*$ is an integral curve of $X_{\tilde{f}}(\alpha)$ we may write:

$$X_{\tilde{f}}(\alpha) \equiv \dot{\alpha} = -ad'_{\nabla\tilde{f}(\alpha)}\alpha \tag{43}$$

This equation, which is to be interpreted as the equation of motion of a Hamiltonian system on $\underline{g}^*$, is often referred to as *Kirillov's equation.* (43) may be rewritten in the form:

$$X_{\tilde{f}}(\alpha) = -ad'_{\nabla\tilde{f}(\alpha)}\alpha = -\frac{d}{dt} Ad'_{\exp t\nabla\tilde{f}(\alpha)}(\alpha)\Big|_{t=o} \tag{44}$$

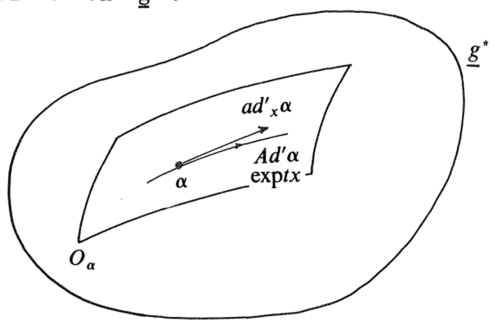
If we make the special choice: $\tilde{f} \equiv \tilde{f}_x$, where:

$$\tilde{f}_x(\alpha) := \alpha(x) \qquad (x\in\underline{g})$$

we get: $\nabla\tilde{f}_x(\alpha) = x$.

Therefore the mapping $\tilde{f}\in C^\infty(\underline{g}^*) \longmapsto \nabla\tilde{f}(\alpha)\in\underline{g}$ is surjective.

Combining this remark with formula (44) we may conclude, that the integral manifolds of the Hamiltonian fields are given by the orbits under the coadjoint action of G on $\underline{g}^*$.



One now easily computes the symplectic form on one of those orbits O_α:

$$\begin{aligned}\omega_\alpha(ad'_x\alpha,\ ad'_y\alpha) &= \omega_\alpha(X_{\tilde{f}_{-x}}(\alpha),\ X_{\tilde{f}_{-y}}(\alpha))\\ &= \{\tilde{f}_{-x},\ \tilde{f}_{-y}\}(\alpha)\\ &= -\alpha([\nabla\tilde{f}_{-x}(\alpha),\ \nabla\tilde{f}_{-y}(\alpha)])\\ &= -\alpha([x,y])\end{aligned} \tag{45}$$

This form is usually referred to as the *Kostant-Kirillov symplectic form* on a coadjoint orbit in the dual of a Lie algebra.

We briefly summarize, what we have done in this paragraph:

- We started with a Poisson manifold $(M, \{ , \})$ as the phase space of a Hamiltonian system, and a symmetry described by a Poisson action $\phi: G\times M \to M$.
- The quotient space M/G was equipped with a Poisson bracket in a natural way.
- The projection $P: M \to M/G$ carries over the solution curves of an invariant Hamiltonian f on M to the solution curves of the associated Hamiltonian $\tilde{f}$ on M/G.
- This general construction was applied to the case: $M = T^*G$, $M/G = \underline{g}^*$.
- The Poisson bracket on $\underline{g}^*$ is given by (42). The integral manifolds of the Hamiltonian fields are the coadjoint orbits of G in $\underline{g}^*$. The symplectic structure on those orbits is given by (45).

§6. BIINVARIANT HAMILTONIAN SYSTEMS ON T^*G

In this section the lines of §§4 and 5 will come together. Central is proposition 6.3, which deals with Poisson commutativity of a certain class of functions on $\underline{g}^*$. This proposition will eventually enable us to prove the involution statement (16) in §8.

Consider a Lie group G, that may be decomposed in two Lie subgroups K and L as in the end of section 4; $G = KL$.
Let $h \in C^\infty(T^*G)$ be left-invariant; $h \circ \bar{L}_g = h \quad \forall g \in G$.

In §4 we saw, that h defines in a natural way a function $\hat{h} \in C^\infty(T^*L)$.
In §5 we saw, that h also defines a function $\tilde{h} \in C^\infty(\underline{g}^*)$;

$$\begin{array}{ccc} h \in C^\infty(T^*G) & \longrightarrow & \hat{h} \in C^\infty(T^*L) \\ \downarrow & & \\ \tilde{h} \in C^\infty(\underline{g}^*) & & \end{array}$$

In this scheme $\hat{h}$ and $\tilde{h}$ may be any functions on T^*L and $\underline{g}^*$ respectively. If however $h \in C^\infty(T^*G)$ is in addition also right-invariant $(h \circ \bar{R}_g = h)$, $\hat{h}$ and $\tilde{h}$ will also possess certain symmetries as will be clear from the following two propositions.

<u>PROPOSITION 6.1</u>. $\hat{h} \in C^\infty(T^*G)$ *right-invariant*

$\Rightarrow$ $h \in C^\infty(T^*L)$ *right-invariant (under the* L*-action on* T^*L *!)*

We will first proof the following lemma.

<u>LEMMA 6.2</u>. $\forall Y \in L$: $\bar{R}_Y$ leaves $J^{-1}(0) \subset T^*G$ invariant and the following diagram commutes:

$$\begin{array}{ccc} J^{-1}(0) \subset T^*G & \xrightarrow{\chi} & T^*L \\ \bar{R}_Y \downarrow & & \downarrow \bar{R}_Y \\ J^{-1}(0) & \xrightarrow{\chi} & T^*L \end{array}$$

PROOF. We recall: $\alpha \in J^{-1}(0) \Longleftrightarrow \alpha(\ker f_*|_{\pi(\alpha)}) = 0$

where: $f: G = KL \to L$ is the projection.

$\forall v \in \ker f_*$: v will be tangent to a curve:

$$\gamma(t) = Z(t)X \qquad Z(t) \subset K, \quad X \in L$$

$$\Rightarrow \quad R_{Y*}v = \frac{d}{dt} Z(t)(XY)$$

and this is again an element of $\ker f_*$ $\forall Y \in L$. Therefore:
$\forall \alpha \in J^{-1}(0)$, $v \in \ker f_*$, $Y \in L$:

$$(\bar{R}_Y\alpha)(v) = \alpha(R_{Y^{-1}*}v) = 0$$

$$\Rightarrow \quad \bar{R}_Y\alpha \in J^{-1}(0) \quad \text{and } \bar{R}_Y \text{ leaves } J^{-1}(0) \text{ invariant.}$$

To prove commutativity of the diagram we observe:

$$\forall \hat{\eta} \in T_{f(\pi(\alpha))Y}L : \exists! \hat{\xi} \in T_{f(\pi(\alpha))}L \quad \text{such that}$$

$$\hat{\xi} = R_{Y^{-1}*}\hat{\eta}$$

$$\forall \hat{\xi} \in T_{f(\pi(\alpha))}L : \exists \xi \in T_{\pi(\alpha)}G \quad \text{such that:}$$

$$\hat{\xi} = f_*\xi$$

We then have:

$$\forall\alpha \in J^{-1}(0) \; : \; (R^*_{\gamma^{-1}} \circ \chi)(\alpha)(\hat{\eta}) = \chi(\alpha)(R_{\gamma^{-1}*}\hat{\eta})$$
$$= \chi(\alpha)(\hat{\xi})$$
$$= \alpha(\xi)$$
$$= (R^*_{\gamma^{-1}}\alpha)(R_{\gamma*}\xi)$$
$$= (\chi \circ R^*_{\gamma^{-1}}\alpha)(f_* \circ R_{\gamma_*}\xi)$$
$$= (\chi \circ R^*_{\gamma^{-1}}\alpha)(R_{\gamma_*} \circ f_*\xi)$$
$$= (\chi \circ R^*_{\gamma^{-1}}\alpha)(\hat{\eta})$$

so: $\quad R^*_{\gamma^{-1}} \circ \chi = \chi \circ R^*_{\gamma^{-1}}$. ∎

PROOF of Proposition 6.1.

$$\hat{h}(R^*_\gamma\hat{\alpha}) = \hat{h}(R^*_\gamma\chi(\alpha))$$
$$= \hat{h}(\chi R^*_\gamma(\alpha))$$
$$= h(R^*_\gamma\alpha)$$
$$= h(\alpha)$$
$$= \hat{h}(\hat{\alpha}).$$ ∎

REMARK. It is clear that if we start from a right-invariant function $\hat{h} \in C^\infty(T^*L)$, the associated function $h := \hat{h}\circ\chi \in C^\infty(T^*G)$ is only right-invariant under the action of L on T^*G.

PROPOSITION 6.3.

$$h \in C^\infty(T^*G) \text{ right-invariant} \overset{i)}{\Longleftrightarrow} \hat{h} \text{ coadjoint invariant}$$
$$\overset{ii)}{\Longleftrightarrow} \operatorname{ad}'_{\nabla\hat{h}(\alpha)}\alpha = 0 \quad \forall\alpha \in \mathfrak{g}^*$$
$$\overset{iii)}{\Longleftrightarrow} \{\tilde{h}, \tilde{g}\}(\alpha) = 0 \quad \forall\tilde{g} \in C^\infty(\mathfrak{g}^*)$$

PROOF.

i) $\alpha \in T^*_{\pi(\alpha)}G\,;\quad R^*_{g^{-1}} : T^*_{\pi(\alpha)}G \to T^*_{\pi(\alpha)g}G$

$$L^*_{\pi(\alpha)g} : T^*_{\pi(\alpha)g}G \to T^*_eG \simeq \underline{g}^*$$

$$\begin{aligned} h(R^*_{g^{-1}}\alpha) &= \tilde{h}(L^*_{\pi(\alpha)g} \circ R^*_{g^{-1}}\alpha) \\ &= \tilde{h}(L^*_g \circ R^*_{g^{-1}} \circ L^*_{\pi(\alpha)}\alpha) \\ &= \tilde{h}(\mathrm{Ad}'_{g^{-1}} L^*_{\pi(\alpha)}\alpha) \qquad (1) \\ h(\alpha) &= \tilde{h}(L^*_{\pi(\alpha)}\alpha) \qquad (2) \end{aligned}$$

from (1) and (2) i) is immediate. ∎

ii) $\tilde{h}$ coadjoint invariant

$$\Longleftrightarrow \frac{d}{dt}\tilde{h}(\mathrm{Ad}'_{\exp t\,x}\alpha) = 0 \qquad \forall x \in \underline{g}\,,\quad \alpha \in \underline{g}^*$$

$$\Longleftrightarrow^* \frac{d}{dt}\tilde{h}(\mathrm{Ad}'_{\exp t\,x}\alpha)\Big|_{t=0} = 0 \qquad \forall x,\alpha$$

$$\Longleftrightarrow \mathrm{ad}'_x\alpha(\nabla\tilde{h}(\alpha)) = 0 \qquad \forall x,\alpha$$

$$\Longleftrightarrow \mathrm{ad}'_{\nabla\tilde{h}(\alpha)}\alpha(x) = 0 \qquad \forall x,\alpha$$

$$\Longleftrightarrow \mathrm{ad}'_{\nabla\tilde{h}(\alpha)}\alpha = 0 \qquad \forall \alpha \in \underline{g}^*. \quad ∎$$

*NB: $\frac{d}{dt}\tilde{h}(\mathrm{Ad}'_{\exp t\,x}\alpha)\Big|_{t=0} = 0 \Rightarrow \frac{d}{dt}\tilde{h}(\mathrm{Ad}'_{\exp t\,x}\alpha)\Big|_{t=s} =$

$$= \frac{d}{dt'}\tilde{h}(\mathrm{Ad}'_{\exp t'\,x}(\mathrm{Ad}'_{\exp s\,x}\alpha))\Big|_{t'=0}$$

$$= \frac{d}{dt'}\tilde{h}(\mathrm{Ad}'_{\exp t'\,x}\beta)\Big|_{t'=0} = 0$$

iii) $\{\tilde{h},\tilde{g}\}(\alpha) = -X_{\tilde{h}}(\alpha)[\tilde{g}]$

$$= \mathrm{ad}'_{\nabla\tilde{h}(\alpha)}\alpha(\nabla\tilde{g}(\alpha))$$

$$= 0 \quad \forall\tilde{g}. \quad ∎$$

We see, that if we start from a biinvariant $h \in C^{\infty}(T^*G)$, we end up with a right-invariant $\hat{h} \in C^{\infty}(T^*L)$. But such $\hat{h}$ defines in a natural way an $\tilde{\tilde{h}} \in C^{\infty}(\underline{\ell}^*)$ by:

$$\forall \alpha \in T^*L: \quad \hat{h}(\alpha) =: \tilde{\tilde{h}}(R^*_{\pi(\alpha)}\alpha)$$

$$(R^*_{\pi(\alpha)}\alpha \in T^*_eL \simeq \underline{\ell}^*)$$

(This is of course completely analogous to the construction in §5, where we started from left-invariant functions on T^*G in stead of right-invariant functions on T^*L) .

The Hamiltonian structure on $\underline{\ell}^*$ is again given by Kirillov's Poisson bracket:

$$\forall \mu \in \underline{\ell}^*: \quad \{\tilde{\tilde{h}}, \tilde{\tilde{g}}\}(\mu) = \mu([\nabla \tilde{\tilde{h}}(\mu), \nabla \tilde{g}(\mu)]) \tag{46}$$

<u>NB</u>: The difference between left- and right-actions is manifest in the sign of the Poisson bracket.

We thus have the following diagram:

$$\begin{array}{ccc} h \in C^{\infty}(T^*G) & \longrightarrow & \hat{h} \in C^{\infty}(T^*L) \\ \text{biinvariant} & & \text{right-invariant} \\ \downarrow & & \downarrow \\ \tilde{h} \in C^{\infty}(\underline{g}^*) & & \tilde{\tilde{h}} \in C^{\infty}(\underline{\ell}^*) \\ \text{coadjoint invariant} & & \end{array}$$

<u>LEMMA 6.4</u>. In the above diagram we have:

$$\tilde{\tilde{h}} \equiv \tilde{h}\Big|_{\underline{\ell}^* \subset \underline{g}^*} \tag{47}$$

PROOF.

$\forall \hat{\alpha} \in \underline{\ell}^* \simeq T_e^* L: \quad \hat{\tilde{\tilde{h}}}(\hat{\alpha}) = \hat{h}(\hat{\alpha})$.

If $f: G = KL \to L$ is again the projection on L, we have:

$$\forall \hat{\xi} \in T_e L: \quad \hat{\xi} = \frac{d}{dt}(X(t))\Big|_{t=0}, \qquad X(t) \subset L, \quad X(0) = e$$

$$f_* \hat{\xi} = \frac{d}{dt} f(X(t))\Big|_{t=0} = \frac{d}{dt} X(t)\Big|_{t=0} = \hat{\xi}.$$

So: $\qquad \hat{\alpha}(\hat{\xi}) = \hat{\alpha}(f_* \hat{\xi}) := \alpha(\hat{\xi})$.

So: $\qquad \chi: \hat{\alpha} \longmapsto \hat{\alpha} \qquad \forall \hat{\alpha} \in T_e^* L$.

We then have:

$$\hat{h}(\hat{\alpha}) = h(\hat{\alpha}) = \tilde{h}(\hat{\alpha}). \quad ∎$$

REMARK. Since $\tilde{\tilde{h}}$ is the restriction of $\tilde{h}$ to $\underline{\ell}^* \subset \underline{g}^*$ and $\tilde{h}$ is coadjoint invariant, one might think, that $\tilde{\tilde{h}}$ is also coadjoint invariant. This is, however, *not* true. We must realize that we are in fact dealing with two distinct actions:

$$Ad^{G'}: G \times \underline{g}^* \to \underline{g}^*$$
$$Ad^{L'}: L \times \underline{\ell}^* \to \underline{\ell}^* .$$

In particular we will see in the example of the next section, that for $X \in L$ the mapping

$$Ad_X^{G'}: \underline{\ell}^* \to \underline{g}^*$$

does not leave invariant $\underline{\ell}^*$, while

$$Ad_X^{L'}: \underline{\ell}^* \to \underline{\ell}^*$$

has this property by definition.

Therefore for $\beta \in \underline{\ell}^*$:

$$\begin{aligned}\tilde{\tilde{h}}(Ad_X^{L'}\beta) &= \tilde{h}(Ad_X^{L'}\beta)\\ &\neq \tilde{h}(Ad_X^{G'}\beta)\\ &= \tilde{h}(\beta) = \tilde{\tilde{h}}(\beta)\end{aligned}$$

At this point we recall Kirillov's equation (43) for the Hamiltonian vector field on $\underline{g}^*$:

$$X_{\tilde{h}}(\alpha) = -ad'_{\nabla\tilde{h}(\alpha)}\alpha \ .$$

Because $\tilde{h}$ is coadjoint invariant, we have using proposition 6.3:

$$X_{\tilde{h}}(\alpha) = 0 \qquad \forall\alpha \in \underline{g}^*$$

meaning that the Hamiltonian systems on coadjoint orbits in $\underline{g}^*$, derived from a biinvariant Hamiltonian system on T^*G, are trivial. This fact will enable us to solve Hamilton's equations on T^*G explicitly. By performing two successive projections from T^*G to T^*L and from T^*L to $\underline{\ell}^*$ we can then solve the Hamiltonian systems on orbits in $\underline{\ell}^*$, which are nontrivial, because $\tilde{\tilde{h}}$ is not coadjoint invariant. One of those Hamiltonian systems in $\underline{\ell}^*$ will appear to be the Toda-lattice, in which we are interested. (see §8)

§7. EXPLICIT CALCULATIONS

In this section we will specify the groups G, K and L and their algebras. By introducing coordinates on these groups, we will be able to derive explicit formulae for the left-actions, (co-)adjoint actions etc.

For G we will choose the general linear group: $G = GL(n,\mathbb{R})$.

PROPOSITION 7.1.

$\forall Q \in GL(n,\mathbb{R})$: $\exists! Z \in O(n,\mathbb{R})$, $\exists! X \in L(n,\mathbb{R})$

such that $Q = ZX$,

where: $O(n,\mathbb{R})$ = *orthogonal group*

$L(n,\mathbb{R})$ = *group of matrices of the form:*

$$\begin{pmatrix} x_{11} & & \phi \\ \vdots & \ddots & \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}, \quad x_{ii} > 0$$

also called: the lower triangular group.

The proof of this proposition is based on the Gramm-Schmidt orthogonalization procedure. It is not difficult but rather technical and therefore we will omit it here.

On $GL(n,\mathbb{R})$ we have a global chart:

$$K: Q \in GL(n,\mathbb{R}) \longmapsto (Q_{ij}) \in \mathbb{R}^{n^2} .$$

This induces global charts on TG and T^*G :

$TG = G \times \underline{g}$ so $X \in T_Q G$: $X = (Q,\dot{Q})$ $\dot{Q} \in \underline{g} = M(n,\mathbb{R})$

$T^*G = G \times \underline{g}^*$ so $\alpha \in T_Q^*G$: $\alpha = (Q,P)$ $P \in \underline{g}^* = M(n,\mathbb{R})$

If $X = \sum_{ij} \dot{Q}_{ij} \frac{\partial}{\partial Q_{ij}}\Big|_Q$ and $\alpha = \sum_{ij} P_{ij}\, dQ_{ij}\Big|_Q$ we have:

$$\alpha(X) = \sum_{ij} P_{ij}\, \dot{Q}_{ij} = \operatorname{tr} P^T \dot{Q}.$$

In these coordinates one now easily derives the following formulae:

i) $$L_{A*}(Q,\dot{Q}) = (AQ,A\dot{Q}) \qquad \forall A,Q \in G\ ,\quad \dot{Q} \in \underline{g} \tag{48i}$$

ii) $$L^*_{A^{-1}}(Q,P) = (AQ,(A^{-1})^T P) \qquad \forall A,Q \in G\ ,\quad P \in \underline{g}^* \tag{48ii}$$

iii) $$\mathrm{Ad}_Q X = QXQ^{-1} \qquad \forall Q \in G\ ,\quad X \in \underline{g} \tag{48iii}$$

$$\mathrm{ad}_X Y = [X,Y] \qquad \forall X,Y \in \underline{g}$$

$$\mathrm{Ad}'_Q P = (Q^{-1})^T P\, Q^T \qquad \forall Q \in G\ ,\quad P \in \underline{g}^*$$

$$\mathrm{ad}'_X P = [P,X^T] \qquad \forall X \in \underline{g}\ ,\quad P \in \underline{g}^*$$

<u>PROOF</u>.

i) trivial

ii) Suppose: $L^*_{A^{-1}}(Q,P) = (AQ,\hat{P})$ then:

$$\begin{aligned} L^*_{A^{-1}}(Q,P)(L_{A*}(Q,\dot{Q})) &= (Q,P)(Q,\dot{Q}) = \operatorname{tr} P^T \dot{Q} \\ &= (AQ,\hat{P})(AQ,A\dot{Q}) \\ &= \operatorname{tr} \hat{P}^T A\dot{Q} = \operatorname{tr}(A^T\hat{P})^T \dot{Q} \qquad \forall \dot{Q} \in \underline{g} \end{aligned}$$

$$\Rightarrow \quad \operatorname{tr} P^T \dot{Q} = \operatorname{tr}(A^T\hat{P})^T\dot{Q} \qquad \forall \dot{Q} \in \underline{g}$$

$$\Rightarrow \quad \hat{P} = (A^{-1})^T P\ . \quad \blacksquare$$

iii) $$\mathrm{Ad}_Q X := L_{Q*} R_{Q^{-1}*} X = QX\,Q^{-1}$$

$$\mathrm{ad}_X Y = [X,Y] \qquad \text{(see definition 5.2)}$$

$$\begin{aligned} \mathrm{Ad}'_Q P(X) &:= P(\mathrm{Ad}_{Q^{-1}} X) \\ &= \operatorname{tr}\, P^T Q^{-1} XQ \\ &= \operatorname{tr}[(Q^{-1})^T P\, Q^T]^T X \qquad \forall X \in \underline{g} \end{aligned}$$

So: $\quad \mathrm{Ad}'_Q P = (Q^{-1})^T P Q^T.$ ∎

$$\mathrm{ad}'_X P(Y) := -P([X,Y])$$
$$= -\mathrm{tr}\, P^T[X,Y]$$
$$= -\mathrm{tr}[P^T,X],Y$$
$$= \mathrm{tr}[P,X^T]^T Y \qquad \forall Y \in \underline{g}$$

So: $\quad \mathrm{ad}'_X P = [P,X^T].$ ∎

In order to get the analogous formulae on TL, T^*L, $\underline{\ell}$ and $\underline{\ell}^*$ we observe, that besides the decomposition of the group G = KL, we also have a decomposition of the algebras:

$$\underline{g} = \underline{k} \oplus \underline{\ell} \tag{49}$$

where $\quad \underline{k}$ = {antisymmetric matrices}

$\underline{\ell}$ = {lower triangular matrices}

define furthermore:

$$\underline{k}^\perp := \{X \in \underline{g}^* \mid \mathrm{tr}\, X^T Y = 0 \quad \forall Y \in \underline{k}\}$$
$$= \{\text{symmetric matrices}\}$$
$$\underline{\ell}^\perp := \{X \in \underline{g}^* \mid \mathrm{tr}\, X^T Y = 0 \quad \forall Y \in \underline{\ell}\}$$
$$= \{\text{strictly}^{**} \text{ upper triangular matrices}\}$$

** i.e.: the diagonal elements are zero.

We then have:

$$\begin{aligned} g^* &= \underline{k}^* \oplus \underline{\ell}^* \\ &= \underline{\ell}^\perp \oplus \underline{k}^\perp \end{aligned} \tag{50a}$$

and this yields the isomorphisms:

$$\underline{k}^* \simeq \underline{\ell}^\perp \quad \text{and} \quad \underline{\ell}^* \simeq \underline{k}^\perp \tag{50^b}$$

We can now introduce coordinates:

$$K: X \in L \longmapsto (X_{11}, X_{21}, X_{22}, X_{31} \dots X_{nn}) \in \mathbb{R}^{\frac{1}{2}n(n+1)}$$

$$TL = L \times \underline{\ell} \quad \text{so:} \quad \xi \in T_X L: \quad \xi = (X, \dot{X}) \quad (\dot{X} \in \underline{\ell})$$

$$T^*L = L \times \underline{\ell}^* \simeq L \times \underline{k}^\perp \quad \text{so:} \quad \hat{\alpha} \in T_X^* L: \quad \hat{\alpha} = (X, R) \quad (R \in \underline{k}^\perp) \ .$$

The decompositions are made explicit by:

$$\forall A \in M(n, \mathbb{R}): \quad A = \Pi_{\underline{k}} A + \Pi_{\underline{\ell}} A$$

and:

$$A = \Pi_{\underline{k}^\perp} A + \Pi_{\underline{\ell}^\perp} A \ ,$$

where:

$$\Pi_{\underline{k}} A = A_+ - (A_+)^T \tag{51^a}$$

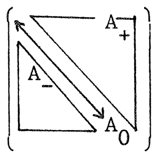
$$\Pi_{\underline{\ell}} A = A_0 + A_- + (A_+)^T \tag{51^b}$$

$$\Pi_{\underline{k}^\perp} A = A_0 + A_- + (A_-)^T \tag{51^c}$$

$$\Pi_{\underline{\ell}^\perp} A = A_+ - (A_-)^T \tag{51^d}$$

and:

A =



These formulae can easily be checked in the case of 2×2 matrices.

We are now ready for the following formulae:

i) $$L_{Y*}(X, \dot{X}) = (YX, Y\dot{X}) \qquad Y, X \in L, \quad \dot{X} \in \underline{\ell} \tag{52i}$$

ii) $L^*_{Y^{-1}}(X,R) = (YX, \Pi_{\underline{k}^\perp}(Y^{-1})^T R)$ $\quad \forall Y,X \in L, \quad R \in \underline{k}^\perp$ (52ii)

iii) $Ad_X Y = XYX^{-1}$ $\quad \forall X,Y \in L$ (52iii)

$ad_X Y = [X,Y]$ $\quad \forall X,Y \in \underline{\ell}$

$Ad'_X R = \Pi_{\underline{k}^\perp}(X^{-1})^T R X^T$ $\quad \forall X \in L, \quad R \in \underline{k}^\perp$

$ad'_X R = \Pi_{\underline{k}^\perp}[R, X^T]$ $\quad \forall X \in \underline{\ell} \quad R \in \underline{k}^\perp$

COMMENT. These formulae only differ from the analogous ones on TG, T^*G etc. in an occasional projection $\Pi_{\underline{k}^\perp}$. We will only proof ii) to show how the projection $\Pi_{\underline{k}^\perp}$ comes in.

Suppose: $L^*_{Y^{-1}}(X,R) = (YX, \hat{R})$ where $X,Y \in L$, $R, \hat{R} \in \underline{k}^\perp$ we then get:

$$(L^*_{Y^{-1}}(X,R))(L_{Y*}(X,\dot{X})) = (X,R)(X,\dot{X})$$

$$= \mathrm{tr}\, R^T \dot{X}$$

$$= (YX,\hat{R})(YX,Y\dot{X})$$

$$= \mathrm{tr}(Y^T\hat{R})^T \dot{X}$$

So: $\quad \mathrm{tr}\, R^T \dot{X} = \mathrm{tr}(Y^T\hat{R})^T \dot{X} \quad \forall \dot{X} \in \underline{\ell}$

$\Rightarrow \quad R = Y^T\hat{R} + Z \quad$ where: $\quad Z \in \underline{\ell}^\perp$

$\Rightarrow \quad \hat{R} = (Y^{-1})^T R - (Y^{-1})^T Z$

(NB: $(Y^{-1})^T Z \in \underline{\ell}^\perp$)

$\hat{R} \in \underline{k}^\perp$ so: $\quad \hat{R} = \Pi_{\underline{k}^\perp}\hat{R} = \Pi_{\underline{k}^\perp}(Y^{-1})^T R$. ∎

(NB: $\Pi_{\underline{k}^\perp}(\underline{\ell}^\perp) = \{0\}$ because $\underline{g}^* = \underline{k}^\perp \oplus \underline{\ell}^\perp$) .

In §6 we constructed the following diagram:

$$
\begin{array}{ccc}
\begin{array}{c} h \in C^{\infty}(T^{*}G) \\ \text{biinv.} \end{array} & \longrightarrow & \begin{array}{c} \hat{h} \in C^{\infty}(T^{*}L) \\ \text{rightinv.} \end{array} \\
\downarrow & & \downarrow \\
\begin{array}{c} \tilde{h} \in C^{\infty}(\underline{g}^{*}) \\ \text{coadj.inv.} \end{array} & & \tilde{\hat{h}} \in C^{\infty}(\underline{\ell}^{*})
\end{array}
$$

where: $\tilde{\hat{h}} \equiv \tilde{h}\Big|_{\underline{\ell}^{*} \subset \underline{g}^{*}}$.

We now mention without proof:

$$\nabla\tilde{\hat{h}}(A) = \Pi_{\underline{\ell}} \nabla \tilde{h}(A) \qquad (A \in \underline{k}^{\perp}) \qquad (53)$$

§8. THE TODA-LATTICE

In this section we will finally fulfill the promises, made in the end of §1. We will get an interpretation of the Lax-equation (14) and the Flaschka-transformation (13). Also we will proof the involution statement (16) and solve the initial value problem.

Consider the situation, that was explained in §7.

$G = KL$ where $G = GL(n,\mathbb{R})$, $K = O(n,\mathbb{R})$, $L = L(n,\mathbb{R})$.

If we start from a left-invariant Hamiltonian $h \in C^\infty(T^*G)$ we have:

$$\begin{array}{ccc} h \in C^\infty(T^*G) & \qquad & (Q,P) \in T^*G \\ \Big\downarrow & & \Big\downarrow L_Q^* \\ \tilde{h} \in C^\infty(\underline{g}^*) & & X := Q^TP \in \underline{g}^* \end{array}$$

$$h(Q,P) = h(L_Q^*(Q,P)) = h(1,Q^TP) =: \tilde{h}(Q^TP) \tag{54}$$

Hamilton's equations on T^*G become:

$$\begin{cases} \dot{Q}_{ij} = \dfrac{\partial h}{\partial P_{ij}} = \dfrac{\partial \tilde{h}(Q^TP)}{\partial P_{ij}} \\[2ex] \dot{P}_{ij} = -\dfrac{\partial h}{\partial Q_{ij}} = -\dfrac{\partial \tilde{h}(Q^TP)}{\partial Q_{ij}} \end{cases}$$

A short calculation yields

$$\begin{cases} \dot{Q} = Q\nabla\tilde{h}\,(Q^TP) \\ \dot{P} \quad P\nabla\tilde{h}^{\,T}(Q^TP) \end{cases} \tag{55}$$

The evolution of $X := Q^TP \in \underline{g}^*$ may now be calculated in two ways:

1) directly: $\dot{X} = \dot{Q}^TP + Q^T\dot{P}$

$$= \nabla\tilde{h}^TQ^TP - Q^TP\nabla\tilde{h}^T$$

$$= [\nabla\tilde{h}^{\,T}(X), X]$$

2) with the aid of Kirillov's equation (43) and formula (48iii):

$$\dot{X} = -\,\mathrm{ad}'_{\nabla\tilde{h}(X)}X = [\nabla\tilde{h}(X)^T, X]\ .$$

In the situation of §7 $h \in C^\infty(T^*G)$ was also right-invariant. We then have according to proposition 6.3:

$$\mathrm{ad}'_{\nabla\tilde{h}(X)}X = 0 \qquad \forall X \in \underline{g}^*$$

and therefore:

$$\dot{X} = 0 \;\Rightarrow\; Q^TP = \text{constant} = Q_o^TP_o\ .$$

The system (55) is now solved trivially:

$$\begin{cases} Q(t) = Q(o)\ \exp(t\,\nabla\tilde{h}(Q^TP)) \\ \\ P(t) = P(o)\ \exp(-t\,\nabla\tilde{h}^T(Q^TP)) \end{cases} \tag{56^a}$$

or equivalently using formula (48ii):

$$(Q(t),\ P(t)) = R^*_{\exp -t\nabla\tilde{h}(Q^TP)}(Q_o, P_o) \tag{56^b}$$

The general reduction scheme was:

$$\begin{array}{ccc} (Q,P) \in J^{-1}(0) \subset T^*G & \xrightarrow{\ \ \chi\ \ } & (X,R) \in T^*L \\ \Big\downarrow L_Q^* & & \Big\downarrow R_X^* \\ X = Q^TP \in \underline{g}^* & & A := \Pi_{\underline{k}^\perp} RX^T \in \underline{\ell}^* \simeq \underline{k}^\perp \end{array}$$

We now calculate the evolution of $A := \Pi_{\underline{k}^\perp} RX^T$ using Kirillov's equation on $\underline{\ell}^*$ (43):

$$\dot{A} = ad'_{\nabla\tilde{\tilde{h}}(A)} A \qquad \text{(-sign because of right-reduction!)}$$

$$= \Pi_{\underline{k}^\perp}[A, (\nabla\tilde{\tilde{h}}(A))^T] \qquad \text{(formula (52iii))}$$

$$= -\Pi_{\underline{k}^\perp}[A^T, \nabla\tilde{\tilde{h}}(A)]^T$$

$$= -\Pi_{\underline{k}^\perp}[A, \nabla\tilde{\tilde{h}}(A)]^T \qquad (A \in \underline{k}^\perp \to A^T = A)$$

$$= -\Pi_{\underline{k}^\perp}[A, \Pi_{\underline{\ell}}\nabla\tilde{h}(A)]^T \qquad \text{(formula (53)}$$

$$= -\Pi_{\underline{k}^\perp}[A, \nabla\tilde{h}(A)]^T + \Pi_{\underline{k}^\perp}[A, \Pi_{\underline{k}}\nabla\tilde{h}(A)]^T$$

$$= \Pi_{\underline{k}^\perp}[A, \Pi_{\underline{k}}\nabla\tilde{h}(A)]^T \qquad ([A, \nabla\tilde{h}(A)] = 0)$$

$$= [A, \Pi_{\underline{k}}\nabla\tilde{h}(A)] .$$

In the last line we used the fact:

$$[\underline{k}^\perp, \underline{k}] \subset \underline{k}^\perp .$$

So we finally get:

$$\dot{A} = [A, \Pi_{\underline{k}}\nabla\tilde{h}(A)] \tag{57}$$

If we want to proceed any further, we will now have to specify the Hamiltonian $\tilde{h} \in C^\infty(\underline{g}^*)$. It is clear that h will have to be a coadjoint invariant function on $\underline{g}^*$. A class of coadjoint invariant functions is given by the so-called *trace-monomials*:

$$\tilde{f}_k(X) := \frac{1}{k}\,\mathrm{tr}\, X^k \qquad k = 1,2\ldots \qquad X \in \underline{g}^*$$

The gradient becomes:

$$\nabla\tilde{f}_k(X) = (X^T)^{k-1} \tag{58}$$

If we choose in particular $\tilde{h} = \tilde{f}_2$, we get for $A \in \underline{k}^\perp$:

$$\nabla\tilde{h}(A) = A$$

equation (57) then becomes:

$$\dot{A} = [A,B] \quad \text{where} \quad B = \Pi_{\underline{k}} A = A_+ - (A_+)^T \tag{59}$$

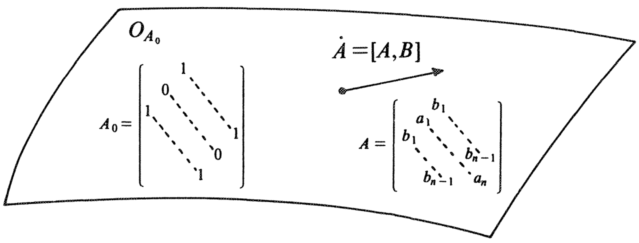
As was explained in §5 the Hamiltonian vector fields on $\underline{\ell}^*$ are tangent to the orbits of the coadjoint action of L on $\underline{\ell}^*$. We now select one very special orbit, namely the one passing through the matrix

$$A_0 := \begin{pmatrix} 0 & 1 & & \emptyset \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & \ddots & 0 \\ \emptyset & & 1 & \end{pmatrix}$$

Using formula (52iii) for the coadjoint action of L on $\underline{\ell}^*$, we can explicitly compute the form of the matrices in $O_{A_0} \equiv \{Ad'_X A_0 \mid X \in L\}$ and get:

$$\forall A \in O_{A_0}: \quad A = \begin{pmatrix} a_1 & b_1 & & \emptyset \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & \ddots & a_n \\ \emptyset & & b_{n-1} & \end{pmatrix} \tag{60}$$

where $\sum_{i=1}^{n} a_i = 0, \quad b_i > 0$



for $B = \Pi_{\underline{k}} A$ we get:

$$B = \begin{pmatrix} 0 & b_1 & & & \\ \cdot & \cdot & \cdot & & \\ -b_1 & & \cdot & \cdot & b_{n-1} \\ & \cdot & & \cdot & \\ & & -b_{n-1} & & 0 \end{pmatrix} \tag{61}$$

If we fill in (60) and (61) in (59) we indeed get the Lax-equation (14) for the Toda-lattice and we may therefore conclude that the Lax-equation (14) is to be interpreted as the equation for the Hamiltonian vector field tangent to the orbit 0_{A_0} .

We must realize, that the set of coordinates a_i, b_i, $\sum a_i = 0$, $b_i > 0$ is just an 'obvious' set of coordinates on the symplectic manifold 0_{A_0}, but they are by no means canonical (≡ Darboux) coordinates.

In fact a rather long calculation yields for the Kostant-Kirillov symplectic form on 0_{A_0} :

$$\omega = \sum_{i=1}^{n-1} \frac{db_i}{b_i} \wedge \left(\sum_{k=1}^{i} da_k\right) \tag{62}$$

If we use $a_1 \ldots a_{n-1}$, $b_1 \ldots b_{n-1}$ as independent coordinates on 0_{A_0} we can consider the coordinate transformation:

$$\begin{cases} a_i = -p_i \\ b_i = e^{q_{i+1}-q_i} \end{cases} \qquad i = 1,2 \ldots n-1 \tag{63}$$

where q_n is defined as: $q_n = -\sum_{i=1}^{n-1} q_i$ (one easily checks that the Jacobian of the transformation $(a_1-a_{n-1}, b_1-b_{n-1}) \to (p_1-p_{n-1}, q_1-q_{n-1})$ is non-zero, so that (63) indeed defines a coordinate transformation.)

If we define furthermore: $p_n = -a_n = -\sum_{i=1}^{n-1} p_i$ we have:

$$\begin{aligned}\omega &= \sum_{i=1}^{n-1} \frac{db_i}{b_i} \wedge \left(\sum_{k=1}^{i} da_k\right) \\ &= -\sum_{i=1}^{n-1} (dq_{i+1}-dq_i) \wedge \left(\sum_{k=1}^{i} dp_k\right) \\ &= \sum_{i=1}^{n-1} dq_i \wedge \left(\sum_{k=1}^{i} dp_k\right) - \sum_{i=2}^{n} dq_i \wedge \left(\sum_{k=1}^{i=1} dp_k\right) \\ &= dq_1 \wedge dp_1 + \sum_{i=2}^{n-1} dq_i \wedge dp_i - dq_n \wedge \sum_{k=1}^{n-1} dp_k \\ &= \sum_{i=1}^{n} dq_i \wedge dp_i \Big|_{CM} \end{aligned} \tag{64}$$

where CM stands for the centre-of-mass manifold in $T^*\mathbb{R}^n$ defined by:

$$\sum_{i=1}^{n} q_i = 0 = \sum_{i=1}^{n} p_i \,.$$

We conclude:

The coadjoint orbit O_{A_0} may be identified with the centre-of-mass manifold CM in $T^*\mathbb{R}^n$. The Flaschka-transformation (13) is to be interpreted as a transformation from a set of obvious coordinates on O_{A_0} to a set of canonical coordinates on CM.

We will now proof the involution statement (16).

<u>PROPOSITION 8.1.</u> *Define*

$$\tilde{\tilde{f}}_k(A) := \frac{1}{k}\,\mathrm{tr}\, A^k \qquad\qquad \forall A \in \underline{k}^\perp \simeq \underline{\ell}^*$$

then

$$\{\tilde{\tilde{f}}_k, \tilde{\tilde{f}}_\ell\}(A) = 0.$$

PROOF. We of course have:

$$\begin{array}{ccc} f_k \in C^\infty(T^*G) & \longrightarrow & \hat{f}_k \in C^\infty(T^*L) \\ \text{biinvariant} & & \text{rightinvariant} \\ \downarrow & & \downarrow \\ \tilde{f}_k \in C^\infty(\underline{g}^*) & & \tilde{\tilde{f}}_k \in C^\infty(\underline{\ell}^*) \\ \text{coadj.invariant} & & \end{array}$$

$$\begin{aligned} \forall A \in \underline{k}^\perp \simeq \underline{\ell}^* : \quad \{\tilde{\tilde{f}}_k, \tilde{\tilde{f}}_\ell\}(A) &:= \{\hat{f}_k, \hat{f}_\ell\}(A) \\ &= \{\hat{f}_k, \hat{f}_\ell\}(\chi(A)) && \text{(see proof lemma 6.4)} \\ &= \{f_k, f_\ell\}(A) && \text{(proposition 4.3iii)} \\ &=: \{\tilde{f}_k, \tilde{f}_\ell\}(A) \\ &= 0 \end{aligned}$$

because $\tilde{f}_k$ is coadjoint invariant. (see proposition 6.3) ∎

We will now briefly sketch how to solve the initial value problem of the Toda-lattice. It should be clear, that once we have a solution curve $(Q(t), P(t))$ for a biinvariant Hamiltonian $h \in C^\infty(T^*G)$ we can perform two successive projections.

$$(Q(t), P(t)) \xmapsto{\chi} (X(t), R(t)) \xmapsto{R_X^*} A(t) = \Pi_{\underline{k}^\perp} RX^T$$

to obtain solution curves $(X(t), R(t))$ of the right-invariant Hamiltonian $\hat{h} \in C^\infty(T^*L)$ and $A(t)$ of the Hamiltonian $\tilde{\tilde{h}} \in C^\infty(\underline{\ell}^*)$. In particular we have at $t = 0$:

$$(Q_o, P_o) \xmapsto{\chi} (X_o, R_o) \xmapsto{R_{X_o}^*} A_0 .$$

In general a lot of different initial conditions on $J^{-1}(0) \subset T^*G$ and on T^*L will yield the same initial condition A_0. In particular, we can choose:

$$(1,A_0) \overset{\chi}{\longmapsto} (1,A_0) \longmapsto A_0 \ .$$

The integral curve passing through $(1,A_0)$ on $J^{-1}(0) \subset T^*G$ is given by (56^a):

$$\begin{cases} Q(t) = Q_o \exp t \nabla \tilde{h}(Q^T P) = \exp t A_0 \\ \\ P(t) = P_o \exp - t \nabla \tilde{h}^T (Q^T P) = A_0 \exp t A_0 \end{cases} \tag{65}$$

$(Q^T P = Q_o^T P_o = A_0 \,, \ \nabla \tilde{h}(Q^T P) = A_0)$

A calculation involving an explicit formula for χ, which we did not display here, yields:

$$A(t) = Ad'_{X(t)} A_0 \tag{66}$$

where $X(t)$ is defined by

$$Q(t) = e^{tA_0} = Z(t)X(t) .$$

We can thus give the following 'receipe' for solving the initial value problem:

1) Give the initial data $q_i(0)$, $p_i(0)$ and construct the matrix:

$$A_0 = \begin{pmatrix} a_1(0) & b_1(0) & & \\ b_1(0) & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1}(0) \\ & & b_{n-1}(0) & a_n(0) \end{pmatrix}$$

2) Perform the ortho-triangular decomposition:

$$e^{tA^0} = Z(t)X(t)$$

3) $$A(t) = Ad'_{X(t)} A_0 = \Pi_{\underline{k}^\perp} [X(t)^{-1}]^T A_0 [X(t)]^T$$

extract the $p_i(t)$'s and $q_i(t)$'s from $A(t)$.

§9. FINAL REMARKS

In this essay we explained the integrability property of the Toda-lattice, by describing it as a Hamiltonian system on a certain coadjoint orbit in the dual of a Lie algebra $\underline{\ell}^*$. The symplectic structure on such an orbit was given by the Konstant-Kirillov symplectic form (45).

There are of course other coadjoint orbits in $\underline{\ell}^*$ and corresponding to them other completely integrable systems. As far as we know, these orbits have not been classified in a systematic way yet. An example of another 'Toda-orbit' and its associated integrable system may be found in an article by Symes [8].

In the introduction of this essay we suggested, that the Toda-lattice may be considered as a discrete analogue of the KdV-equation. In fact it is known that the KdV-equation can be written as a Hamiltonian system, possessing a denumerable sequence of constants of the motion, which are in involution with respect to a certain Poisson bracket. (See Gardner et al. [9])

Adler [10], inspired by the results of Gel'fand and Dikii [11], showed that the Toda-lattice and the KdV-equation may be treated in an analogous way. He described the KdV-equation as a completely integrable Hamiltonian system on a coadjoint orbit in a certain infinite dimensional Lie algebra, namely the symbol algebra of pseudo-differential operators. The symplectic structure on the infinite dimensional orbit is again provided by the Kostant-Kirillov form. (generalized to the ∞-dimensional case.) The associated Poisson bracket coincides with Gardner's bracket mentioned above.

More recently Flaschka [12] et al. demonstrated that the AKNS-system of evolution equations, containing all known soliton equations such as Korteweg-de Vries, Sine-Gordon and nonlinear Schrödinger-equation, may be described as a set of Hamiltonian systems on a coadjoint orbit of another infinite dimensional Lie algebra, namely the Kac-Moody algebra $s\ell(2,\mathbb{C}) \otimes \mathbb{C}(\xi,\xi^{-1})$.

With these short remarks we hope to have given the reader an impression of the importance of the Kostant-Kirillov symplectic structure for describing discrete and continuous integrable systems.

REFERENCES

[1] TODA, M.: *Theory of nonlinear lattices* (Springer series in solid state sciences 20(1981)).

[2] OLIVE, D.I. - TUROK, N.: *Algebraic structure of Toda systems*, Nuclear Physics B220 [FS8] (1983).

[3] FLASCHKA, H.: Physical Review B9 (1974), p.1924.

[3] FLASCHKA, H.: Progress of theoretical physics 51(1974), p.703.

[4] LAX, P.: *Integrals of nonlinear equations of evolution and solitary waves*. Comm. Pure Appl.Math. 21(1968).

[5] ARNOLD, V.I.: *Mathematical methods of classical mechanics* (Springer Verlag, New York 1978).

[6] KIRILLOV, A.: *Local Lie algebras*, Russian mathematical surveys 31(1976).

[7] WEINSTEIN, A.: *The local structure of Poisson manifolds*, Journal of differential geometry 18(1983).

[8] SYMES, W.W.: *Hamiltonian group actions and integrable systems*, Physica 1D (1980).

[9] GARDNER - GREENE - KRUSKAL - MIURA: *Korteweg-de Vries equation and generalizations* VI, *methods of exact solution*, Comm. Pure Appl. Math. 27(1974).

[10] ADLER, M.: *On a trace functional for formal pseudo-differential operators and the symplectic structure of the Korteweg-de Vries type equations*, Invent.Math. 50(1979).

[11] GEL'FAND - DIKII: *Fractional powers of operators and Hamiltonian systems*, Functional analysis and its applications, 10(1976).

[12] FLASCHKA - NEWELL - RATIU: *Kac-Moody Lie algebras and soliton equations*, Physica 9D (1983).

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek. H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.

CWI SYLLABI

1 Vacantiecursus 1984: *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981-1982. Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg, et al. *Testing statistical hypotheses: worked solutions.* 1984.

4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.

5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.

6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982-1983. Mathematical structures in field theories.* 1985.

7 Vacantiecurus 1985: *Variatierekening.* 1985.

8 G.M. Tuynman. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.1 Geometric quantization.* 1985.

9 J. van Leeuwen, J.K. Lenstra (eds.). *Parallel computers and computations.* 1985.

10 Vacantiecursus 1986: *Matrices.* 1986.

11 P.W.H. Lemmens. *Discrete wiskunde: tellen, grafen, spelen en codes.* 1986.

12 J. van de Lune. *An introduction to Tauberian theory: from Tauber to Wiener.* 1986.

13 G.M. Tuynman, M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.2.* 1987.

14 Vacantiecursus 1987: *De personal computer en de wiskunde op school.* 1987.