Poisson Brackets (An Unpedagogical Lecture)

RES JOST

Seminar für Theoretische Physik der ETH, Zürich, Switzerland

INTRODUCTION

It was P.A.M. Dirac who emphasized most strongly the significance of the *Poisson brackets* in classical analytical dynamics [see, e.g., P.A.M. Dirac, *Quantum Mechanics* (Clarendon Press, Oxford, 1947), 3rd ed., Sec. 21, p. 84 and earlier editions]. He not only did this but he also gave what turns out to be a complete axiomatic characterization of the Poisson brackets [see Dirac, Eqs. (2)–(6); see also our Sec. 4 (a–f)]. It appears all the more astonishing that the first discussion of these axioms (by a physicist at least) seems to be one by W. Pauli [Nuovo Cimento **10**, 648–667 (1953), Sec. 2]. He shows that, locally at least, Diracs axioms lead to a phase space with coordinates and canonically conjugate momenta (see theorem 4 in Sec. 4 below).

Once this is established, the "correct procedure" consists of course in a complete elimination of any kind of coordinates. The principle results of general analytical dynamics have to be derived by the exclusive use of Poisson brackets. This is what I propose to do in this note. As it turns out the physicists have "always" implicitly used the modern definition of vector fields (see Sec. 1 below) as (special) operators on a suitable function space.

Our enterprise is of course only an exercise of doubtful physical significance. It shares, however, the meager physical content with analytical dynamics in general.[1]

If the present note should serve any purpose it has to be rather broad and has to include much material which can be found in books. This is bad enough. What makes the situation worse is the fact that the writer, being pressed by other obligations, was unable to look up the relevant literature. Everything he says may have been published already in much better form.

The work may, however, serve as a suggestion to all those who are under the obligation to teach analytical dynamics regularly and feel rather tired of the many p's and q's which they have to write on the blackboard.

It is a pleasure to me to dedicate this modest note to the great American teacher of theoretical physics, J. Robert Oppenheimer.

[1] It is nevertheless true that historically this purely formal structure played a well known and crucial part in the development of quantum mechanics. This should serve as a warning to all those who declare any kind of purely formal development *a priori* as "unphysical." Things are not that simple!

1. DIFFERENTIABLE MANIFOLDS, MAPPINGS, ONE PARAMETER GROUPS, VECTORS AND VECTOR FIELDS[2]

(A) In order to avoid arguments of differentiability we restrict the discussion to differentiable manifolds. A differentiable manifold is a pair $(V,\mathcal{E})$ of a connected Hausdorff space with points $p,q,\cdots$ which has a countable basis of open sets and a family $\mathcal{E}$ of continuous real-valued functions $f,g,\cdots$ over V. $\mathcal{E}$ should satisfy the following conditions:

(a) If a function h coincides in a suitable neighborhood $N(p)$ of every point $p \in V$ with an element of $\mathcal{E}$, then h is itself an element of $\mathcal{E}$.

(b) If F is a C^∞ function over R^k and if $f_1, f_2, \cdots, f_k \in \mathcal{E}$ then also $F(f_1, f_2, \cdots, f_k) \subset \mathcal{E}$.

(c) In every point $p \in V$ exists a neighborhood $N(p)$ and n elements $f^1, f^2, \cdots, f^n \in \mathcal{E}$ such that $x^k = f^k(q)$ is a homomorphic mapping of $N(p)$ on an open set in R^n. Every element of $\mathcal{E}$ restricted to $N(p)$ agrees with $F(f^1, f^2, \cdots, f^n)$ for a suitable choice of the C^∞ function F over R^n.

Remarks: (1) The number n (which is independent of p) is the dimension of $(V, \mathcal{E})$. (2) According to (b) it is no restriction to assume that $f^k(p) = 0$. For suitable $\alpha > 0$ the sets

$$\{q|\ |f^k(q)| < \alpha, k = 1,2,\cdots,n\}$$

from a basis of "cubic" neighborhoods of p. (3) $x^1, x^2, \cdots, x^n$ are *local coordinates* in the point p. (4) $\mathcal{E}$ forms a ring with respect to ordinary addition and multiplication of functions. $\mathcal{E}$ contains the constants $\{c\}$ as a subring. (5) The *support* of a function f, supp f, is the smallest closed set in V outside of which f vanishes.

The following lemma is a useful tool for later work.

Lemma 1. Let $x^1, x^2, \cdots, x^n$ be local coordinates in the point p. Every function $f(q) = \hat{f}(x)$ allows for $|x^k| < \alpha$ the expansion

$$\hat{f}(x) = f(p) + \sum_{k=1}^{n} x^k \hat{f}_{,k}(0) + \tfrac{1}{2} \sum_{k,l=1}^{n} x^k x^l A_{kl}(x) \quad (1.1)$$

[2] For this and the next section see e.g. K. Nomizu, *Lie Groups and Differential Geometry* (Mathematical Society of Japan, Tokyo 1956), Chapter I.

with C^∞ functions A_{kl}. $\hat{f}_{,k}(x)$ stands for $\partial\hat{f}(x)/\partial x^k$.

Proof: Clearly

$$\hat{f}(x) = f(p) + \int_0^1 dt \sum_k x^k \hat{f}_{,k}(tx)dt$$
$$= f(p) + \sum_k x^k A_k(x) ,$$

where $A_k(x) = {}_0\!\int^1 \hat{f}_{,k}(tx)dt$ is a C^∞ function. Repeating the argument for A_k leads to the result (1).

(B) A mapping Φ of V into V' induces a mapping of the functions $\mathcal{E}'$ over V' into functions over V by the following convention:

$$f(p) = f'(p') \quad \text{for} \quad p' = \Phi(p) .$$

We write $f = f' \circ \Phi$. The mapping Φ is *differentiable*, if $f' \circ \Phi \in \mathcal{E}$ whenever $f' \in \mathcal{E}'$. Clearly we have $(f' + g') \circ \Phi = f' \circ \Phi + g' \circ \Phi$ and $(f'g') \circ \Phi = (f' \circ \Phi)(g' \circ \Phi)$. The constants are mapped into the same constants. We restrict ourselves to differentiable mappings.

A (differentiable) *transformation* Φ is a one-to-one (differentiable) mapping of V onto V of which the inverse is also differentiable.

A differentiable one-parameter group is a family of transformations Φ_t, $t \in R$ which satisfies

(1a) $V \times R \xrightarrow{\Phi t} V$ is a differentiable mapping,

(1b) $\Phi_{t+s} = \Phi_t \Phi_s$.

According to (a) $f_t = f \circ \Phi_t$, $f \in \mathcal{E}$, is C^∞ in t. If we write for the derivative $\dot{f}_t = df_t/dt$ at $t = 0$, $\dot{f}_0 = L(f)$ then $L(f)$ has the following properties:

(2 a) $L(f) \in \mathcal{E}$ for $f \in \mathcal{E}$,

(2 b) $L(f + g) = L(f) + L(g)$,

(2 c) $L(f \cdot g) = f \cdot L(g) + L(f) \cdot g$,

(2 d) $L(c) = 0$.

The operator L which maps $\mathcal{E}$ into $\mathcal{E}$ is the *infinitesimal generator* of Φ_t.

(C) An operator L which satisfies [B, (2 a)–(2 d)] is by definition a *contravariant vector field*. All contravariant vector fields form a *linear space* $\mathfrak{L}$ *over the ring* $\mathcal{E}$. Scalar multiplication is defined by $(gL)(f) = g \cdot L(f)$.

Every one parameter group Φ_t defines a contravariant vector field, its infinitesimal generator. The converse is however not true (see Appendix). A group Φ_t however is uniquely defined by its infinitesimal generator.

We say that L vanishes in p, if $L(f)(p) = 0$ for all $f \in \mathcal{E}$. The support of L, supp L is the smallest closed set outside of which L vanishes.

Lemma 2. If $f(q) = g(q)$ for $q \in N(p)$ then $L(f)(p) = L(g)(p)$. Therefore, in view of (2 b), supp $L(f) \subset$ supp $f \cap$ supp L.

Proof. Due to (2 b) it is sufficient to show that $f(q) = 0$ for $q \in N(p)$ has the consequence that $L(f)\ (p) = 0$. If f vanishes in $N(p)$ then $N'(p) \subset N(p)$ and $g \in \mathcal{E}$ exist such that (1) $g(q) = 1$ for $q \in N'(p)$; (2) $f \cdot g = 0$. Now (2 c) and (2 d) imply $0 = L\ (fg) = L\ (f) \cdot g + f \cdot L(g)$, but in $N'(p)$ the second term vanishes and the first term on the right hand side equals $L(f)$.

According to lemma 2 the value of $L(f)\ (q)$ in a neighborhood $N(p)$ depends only on the restriction of f to this neighborhood. If a cubic neighborhood $N(p)$ is chosen, we can use local coordinates and the representation (1) of f [restricted to $N\ (p)$]. At the point p we then find [using (2 b), (2 c), (2 d)]

$$L(f)(p) = \sum_k L(x^k)(p) \cdot \hat{f}_{,k}(0) = \sum_k a^k \hat{f}_{,k}(0) ;$$

at a general point $q \in N(p)$

$$L(f)(q) = \sum_k a^k(x) \hat{f}_{,k}(x) , \quad q \in N(p) . \qquad (1.2)$$

This leads us back to the familiar definition of a contravariant vector field. The functions a^k are C^∞ functions according to (2 a).

For $L(f)\ (p)$ we also write $L_p(f)$. This we do for the following reason: L_p is a functional over $\mathcal{E}$ with values in R. All functionals $\{L_p\}$ form an n-dimensional vector space T_p. T_p is, by definition, the *tangent space of V at the point p*. L_p is completely characterized by the properties

(3 a) $L_p(f + g) = L_p(f) + L_p(g)$,

(3 b) $L_p(fg) = f(p)L_p(g) + L_p(f)g(p)$,

(3 c) $L_p(c) = 0$.

A vector field L attributes to every point p an element of T_p in a differentiable manner.

(D) Let Φ be a mapping of V into V' and let $p' = \Phi(p)$. We now form

$$L_p(f' \circ \Phi) = L'_{p'}(f') . \qquad (1.3)$$

A trivial verification shows that $L'_{p'}$, satisfies (3 a)–(3 c) and therefore is an element of the tangent space $T'_{p'}$, of V' at p'. The mapping $L_p \to L'_{p'}$, is *linear*. This linear transformation is denoted by $\Phi'(p)$ and is the *differential* of Φ at the point p:

$$L'_{p'} = \Phi'(p) L_p . \qquad (1.4)$$

(E) From two vector fields L_1 and L_2 we can construct a third vector field by

$$[L_1,L_2](f) = L_1(L_2(f)) - L_2(L_1(f))\,. \quad (1.5)$$

The verification that $[L_1, L_2](f)$ satisfies (2 a)–(2 d) is again trivial as is the verification of the *Jacobi identity:*

$$[L_1,[L_2,L_3]] + [L_2,[L_1,L_3]] + [L_3,[L_1,L_2]] = 0\,. \quad (1.6)$$

The space $\mathfrak{L}$ of all contravariant vector fields therefore forms a Lie ring.

2. COVARIANT VECTOR FIELDS AND TENSOR FIELDS

A linear form ω over $\mathfrak{L}$ with values in $\mathcal{E}$ is by definition a *covariant vector field.* $\omega(L)$ satisfies

(4 a) $\omega(L) \in \mathcal{E}$,

(4 b) $\omega(L_1 + L_2) = \omega(L_1) + \omega(L_2)$,

(4 c) $\omega(gL) = g\omega(L)$.

The definition of the support of ω is left to the reader, as is the formula

$$\text{supp } \omega(L) \subset \text{supp } \omega \cap \text{supp } L.$$

What we are going to prove, however is

Lemma 3. The value $\omega(L)$ (p) depends only on L_p. This dependence is linear.

Proof. It is sufficient to show that $\omega(L)$ $(p) = 0$ if $L_p = 0$. Since $\omega(L)$ (p) depends only on the local behavior of L, we assume that L has its support in a cubic neighborhood of p to local coordinates x. According to (1.2)

$$L(f)(q) = \textstyle\sum_k a^k(x)\hat{f}_{,k}(x)$$

and, due to $L_p = 0$, $a^k(0) = 0$. In addition supp $a^k \subset \{x \mid |x^k| < \alpha\}$. Now lemma 1 tells us that

$$a^k(x) = \textstyle\sum_l A^k_l(x)\cdot x^l$$

with supp $A^k_l \subset$ supp a^k.

$$L(f)(q) = \textstyle\sum_l x^l \sum_k A^k_l(x)\hat{f}_{,}f_k(x)$$

and finally

$$L(f)(q) = \textstyle\sum_l x^l(q)L_l(f) \quad \text{or} \quad L = \sum_l x^l(q)L_l\,,$$

where

$$x^l(p) = 0 \quad \text{and} \quad L_l(f)(q) = \textstyle\sum_k A^u_l(x)\hat{f}_{,k}(x)\,.$$

According to (4 b), (4 c) $\omega(L)$ $(p) = 0$.

We can therefore write

$$\omega(L)(p) = \omega_p(L_p)\,, \quad (2.1)$$

where $\omega_p(L_p)$ is a linear form over the tangent space $T_p.\omega_p$ is therefore an element of the dual space T^*_p. A covariant vector field attributes to every point p an element ω_p from T^*_p.

The covariant vector fields form again a linear space $\mathfrak{W}$ over $\mathcal{E}$. The definition of the linear operations is evident.

For fixed f, $L(f)$ is itself a linear functional over $\mathfrak{L}$ with values in $\mathcal{E}$ and thus defines a covariant vector field df by

$$df(L) = L(f)\,. \quad (2.2)$$

Not every covariant vector field is a *differential* df of a function. Those which are, are called *exact.* The following rules are consequences of (2 a)–(2 d):

(5 a) $d(f + g) = df + dg$,

(5 b) $d(f\cdot g) = fdg + gdf$,

(5 c) $dc = 0$.

A linear form λ over $\mathfrak{W}$ with values in $\mathcal{E}$ determines uniquely a contravariant vector field L by the values it takes for exact differentials. In view of (5 a), (5 b), (5 c)

$$L(f) = \lambda(df) \quad (2.3)$$

satisfies (2 a)–(2 d). On the other hand we find that every contravariant vector field L determines uniquely a linear form λ by linear extension of (2.3). This is the case because locally $dx^1, dx^2,\cdots,dx^n$ form a basis of $\mathfrak{W}$ restricted to a neighborhood of p. In future we will use L and λ interchangeably.

A skew-symmetric contravariant tensor field Λ_k of rank k is a skew-symmetric multilinear form over $\mathfrak{W}$ with values $\Lambda_k(\omega_1,\omega_2,\cdots,\omega_k) \in \mathcal{E}$. We are specially interested in bilinear forms $\Lambda(\omega_1,\omega_2)$. Restricted to exact differentials, they define a mapping of $\mathcal{E} \times \mathcal{E}$ into $\mathcal{E}$:

$$\Lambda(df,dg) = \{f,g\} \quad (2.4)$$

with the properties

(6 a) $\{f,g\} = -\{g,f\} \in \mathcal{E}$,

(6 b) $\{f,g_1 + g_2\} = \{f_1g_1\} + \{f_1g_2\}$,

(6 c) $\{f,g_1g_2\} = \{f_1g_1\}g_2 + g_1\{f_1g_2\}$,

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

The properties (6 a)–(6 d) are characteristic for the restriction of a skew-symmetric contravariant tensor field to exact differentials. This tensor field is uniquely defined by (2.4).

In local coordinates we obtain, exactly by the steps which led us to (1.2), for $\{f_1 g\}$ the expression

$$\{f,g\}(q) = \sum_{k,l} \eta^{kl}(x)\hat{f}_{,k}(x)\hat{g}_{l}(x)$$

with C^∞ functions $\eta^{kl} = -\eta^{lk} = \{x^k,x^l\}$.

A skew-symmetric covariant tensor field Ω_k of rank k is a skew-symmetric multilinear form over $\mathfrak{L}$ with values $\Omega_k(\lambda_1,\lambda_2,\cdots,\lambda_k) = \Omega_k(L_1,L_2,\cdots,L_k) \in \mathcal{E}$. Again we are mainly interested in skew-symmetric tensor fields of rank 2.

Every covariant vector field gives rise to such a tensor field by the equation

$$(d\omega)(L_1,L_2) = L_1(\omega(L_2)) - L_2(\omega(L_1)) - \omega([L_1,L_2]) \ . \tag{2.5}$$

The verification is again elementary. $d\omega$ is the *differential of* ω. The definitions of $[L_1,L_2]$ and of df (2) lead to

$$d(df) = 0 \ . \tag{2.6}$$

A covariant field ω for which $d\omega = 0$ is *closed.* According to (2.6) every exact field is closed. It is well known that the converse is in general not true. Locally however, in a simply connected neighborhood $N(p)$ of any point p, $d\omega = 0$ implies $\omega = dF$, where F is a C^∞ function in $N(p)$.

Similarly we define the differential of a skew-symmetric tensor field Ω of rank 2 by

$$\begin{aligned}(d\Omega)(L_1,L_2,L_3) &= L_1(\Omega(L_2,L_3)) + L_2(\Omega(L_3,L_1)) \\ &+ L_3(\Omega(L_1,L_2)) - \Omega([L_1,L_2],L_3) \\ &- \Omega([L_2,L_3],L_1) - \Omega([L_3,L_1],L_2) \ . \end{aligned} \tag{2.7}$$

Ω is evidently skew-symmetric and satisfies $d\Omega(c,L_2,L_3) = 0$. The verification that $d\Omega$ is multilinear is again purely a matter of routine. Similarly as above we find (it is sufficient to check this for $\omega = gdf$)

$$d(d\omega) = 0 \ . \tag{2.8}$$

The tensor fields $\Omega = d\omega$ are *exact*, the tensor fields for which $d\Omega = 0$ are *closed*. Every exact field is closed but the converse is again in general only locally true, e.g., in a cubic neighborhood. If the *second Betti number* of V vanishes then every closed tensor field of rank 2 is also exact.

3. MANIFOLDS WITH SYMPLECTIC STRUCTURE

A contravariant skew-symmetric tensor field $\Lambda(\omega_1,\omega_2)$ induces a *symplectic structure* into every cotangent space T_p^*. In fact $\Lambda(\omega_1,\omega_2)(p) = \Lambda_p(\omega_{1,p},\omega_{2,p})$ generates a antisymmetric bilinear form in T_p^*. The form Λ_p is *nondegenerate*, if $\Lambda_p(\omega_1,\omega_2) = 0$ for all ω implies $\omega_1 = 0$. The field Λ is *nowhere degenerate,* if Λ_p is nondegenerate for all p. This is only possible, if the dimension of V is even: $n = 2f, f = 1,2,3\cdots$ and if V is *orientable* (see Sec. 5). These conditions are only necessary.

We say that $(V,\mathcal{E})$ carries a *symplectic structure*, if it carries a nowhere singular skew-symmetric contravariant tensor field Λ of rank 2. Such a manifold may be denoted by $(V,\mathcal{E},\Lambda)$.

As in Sec. 2 we introduce the notation $\{f,g\} = \Lambda(df,dg)$. The properties of this bracket symbol are given by (6 a)–(6 d).

Since Λ is nowhere singular it maps $\mathfrak{W}$ *one to one onto* $\mathfrak{L}$ by the equations

$$\lambda_\omega(\omega_1) = \Lambda(\omega,\omega_1) \ . \tag{3.1}$$

If we express $\Lambda(\omega_1,\omega_2)$ in terms of $\lambda_{\omega 1},\lambda_{\omega 2}$ by

$$\Lambda(\omega_1,\omega_2) = \Omega(\lambda_{\omega 1},\lambda_{\omega 2}) \ , \tag{3.2}$$

then Ω is a covariant skew-symmetric nowhere singular tensor field. The inversion of (1) is given by

$$\omega_\lambda(\lambda_1) = \Omega(\lambda_1,\lambda) \ . \tag{3.3}$$

A manifold with symplectic structure can clearly also be defined as a manifold which carries a skew-symmetric nowhere singular covariant tensor field of rank 2.

For λ_{df} we also write L_f and we have

$$L_f(g) = \{f,g\} \ . \tag{3.4}$$

In words: $\mathcal{E}$ is mapped into $\mathfrak{L}$. Of special interest to us is the commutator of L_f and L_g

$$[L_f,L_g](h) = L_{\{f,g\}}(h) + J(f,g,h) \tag{3.5}$$

where

$$J(f,g,h) = \{f\{g,h\}\} + \{g\{h,f\}\} + \{h\{f,g\}\} \ . \tag{3.6}$$

A purely mechanical verification shows, that $J(f,g,h)$ satisfies the conditions

(a) $J(f,g,h) \in \mathcal{E}$ is completely antisymmetric in f,g,h ,

(b) $J(f_1 + f_2,g,h) = J(f_1,g,h) + J(f_2,g,h)$,

(c) $J(f_1f_2,g,h) = f_1J(f_2,g,h) + J(f_1,g,h)\cdot f_2$,

(d) $J(c,g,h) = 0$.

J determines therefore uniquely an antisymmetric, contravariant tensor field of rank 3 by the equation

$$\Lambda_3(df,dg,dh) = J(f,g,h) \ . \tag{3.7}$$

There is an intimate connection between $J(f,g,h)$ and $d\Omega$. According to (2.7) we have

$$\begin{aligned} d\Omega(L_f,L_g,L_h) &= L_f(\Omega(L_g,L_h)) - \Omega([L_f,L_g],L_h) + \text{cycl.} \\ &= L_f(\{g,h\}) - [L_f,L_g](h) + \text{cycl.} \end{aligned}$$

This equals, due to (3.5) and (3.6),

$$d\Omega(L_f,L_g,L_h) = [\{f,\{g,h\}\} - \{\{f,g\},h\} + \text{cycl.}] - 3J(f,g,h)\,,$$

or, finally,

$$d\Omega(L_f,L_g,L_h) = -J(f,g,h)\,. \tag{3.8}$$

Up to a sign therefore Λ_3 is the contravariant form of $d\Omega$.

In order to motivate the following argument we introduce the one-parameter group Φ_t with the infinitesimal generator L. What we want to compute is the change of $\{f,g\}$ under Φ_t:

$$(d/dt)[\{f\circ\Phi_t,g\circ\Phi_t\} - \{f,g\}\circ\Phi_t]|_{t=0}\,.$$

This quantity equals

$$\Delta L(f,g) = \{L(f),g\} + \{f,L(g)\} - L(\{f,g\})\,. \tag{3.9}$$

We use now (9) as definition for ΔL for an arbitrary vector field L. A simple verification shows that ΔL is an antisymmetric contravariant tensor field: it satisfies conditions analogous to (a)–(d) above.

The covariant form of ΔL is related to $d\omega_L$. In order to find this relation we specialize L first to an expression $aL_h = \lambda_{a\cdot dh}$. The general expression is then obtained by summation. Equation (3.9) leads to

$$\begin{aligned}\Delta(aL_h)(f,g) &= \{a\{h,f\},g\} + \{f,a\{h,g\}\} - a\{h,\{f,g\}\}\\ &= \{a,g\}\{h,f\} + \{f,a\}\{h,g\} - aJ(f,g,h)\\ &= -\{f,a\}\{g,h\} + \{g,a\}\{f,h\}\\ &\quad - \Lambda_3(df,dg,a\cdot dh)\,.\end{aligned} \tag{3.10}$$

On the other hand

$$\begin{aligned}d(adh)(L_f,L_g) &= L_f(a)L_g(h) - L_f(h)L_g(a)\\ &= \{f,a\}\{g,h\} - \{f,h\}\{g,a\}\end{aligned} \tag{3.11}$$

such that

$$\Delta(aL_h) = -d(adh)(L_f,L_g) + \Lambda_3(df,dg,adh) \tag{3.12}$$

and in general

$$\Delta L(f,g) = -d\omega_L(L_f,L_g) + \Lambda_3(\omega_L,df,dg) \tag{3.13}$$

or, alternatively,

$$\Delta L(f,g) = -d\omega_L(L_f,L_g) - d\Omega(L,L_f,L_g)\,. \tag{3.14}$$

We close this section with a general formula which contains (3.13) and (3.14) as special cases:

$$\begin{aligned}&\Lambda(\omega_1,d\Lambda(\omega_2,\omega_3)) + \Lambda(\omega_2,d\Lambda(\omega_3,\omega_1)) + \Lambda(\omega_2,d\Lambda(\omega_1,\omega_2))\\ &\quad - \Lambda_3(\omega_1,\omega_2,\omega_3) + d\omega_1(\lambda_{\omega2},\lambda_{\omega3}) + d\omega_2(\lambda_{\omega3},\lambda_{\omega1})\\ &\qquad + d\omega_3(\lambda_{\omega1},\lambda_{\omega2}) = 0\,.\end{aligned} \tag{3.15}$$

The correctness of (3.15) can be seen as follows: firstly (3.15) is correct for $\omega_k = df_k$ because it agrees then with (3.6). Secondly the expression (3.15) is additive and antisymmetric in all arguments ω_k. Thirdly, if (3.15) is correct for ω_1, ω_2, ω_3 it is also correct for $g\omega_1$, ω_2, ω_3. Therefore (3.15) is generally correct.

4. CANONICAL MANIFOLDS

At last we come to the manifolds in which we are actually interested. These manifolds are generalizations of the classical phase space of analytical dynamics.

Definition. A manifold $(V,\mathcal{E})$ is canonical, if a bracket $\{f,g\}$ (Poisson bracket) with the following properties exists:

(a) $\{f,g\} = -\{g,f\} \in \mathcal{E}$;

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

(c) $\{f_1f_2,g\} = f_1\{f_2,g\} + \{f_1,g\}f_2$;

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

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

(f) The tensor field defined by $\Lambda(df,dg) = \{f,g\}$ is nowhere degenerate.

Remark. A canonical manifold has, according to (a)–(d), (f) a symplectic structure. This symplectic structure is restricted by the condition $J(f,g,h) = 0$. This restriction is the *Jacobi identity* (e). The dimension of V is even: $n = 2f$ and V is orientable (see Sec. 5).

Equation (3.8) immediately leads to

Theorem 1. The covariant form $\Omega(L_1,L_2)$ of $\Lambda(\omega_1,\omega_2)$ is closed:

$$d\Omega = 0\,. \tag{4.1}$$

A canonical structure is therefore equivalently characterized by a covariant, skew-symmetric, nowhere-singular closed tensor field Ω of rank 2.

We will not give a complete proof of the following:

Theorem 2 (Pauli). To every point p there exist local canonical coordinates $p_1,\cdots,p_f$, $q_1,\cdots,q_f$ such that

$$\{f,g\}(q) = \sum_k \left(\frac{\partial \hat{f}}{\partial p_k}\frac{\partial \hat{g}}{\partial q_k} - \frac{\partial \hat{f}}{\partial q_k}\frac{\partial \hat{g}}{\partial p_k}\right) \tag{4.2}$$

for $q \in N(p)$.

Proof. Since $d\Omega = 0$,Ω is locally exact:

$$\Omega = d\eta\,,\quad \eta = \textstyle\sum_k a_k dx^k\,. \tag{4.3}$$

The classical reduction procedure[3] of the Pfaffian form (4.3) leads to the normal form

$$\eta = \sum_{k=1}^{f} p_k dq_k . \tag{4.4}$$

(4.4) leads immediately to (4.2).

We know from Sec. 3 that $\mathcal{E}$ is mapped into $\mathfrak{L}$ by

$$L_f(g) = \{f,g\} . \tag{4.5}$$

The Poisson bracket $\{f,g\}$ induces in $\mathcal{E}$ the structure of a Lie ring. Equation (3.5) now leads to

Theorem 3. The mapping $f \rightarrow L_f$ maps $\mathcal{E}$ (as a Lie ring) homomorphically into a subring of $\mathfrak{L}$:

$$[L_f,L_g] = L_{\{f,g\}} . \tag{4.6}$$

Theorem 3 allows a slight but not unimportant generalization. In order to prepare for this we remind the reader, that $\Lambda(\omega_1,\omega_2)$ maps $\mathfrak{W}$ onto $\mathfrak{L}$ by

$$\lambda_\omega(\omega_1) = \Lambda(\omega,\omega_1) \tag{4.7}$$

Instead of (4.7) we also write (risking some confusion however)

$$L_\omega(g) = \Lambda(\omega,dg) . \tag{4.7'}$$

The confusion consists in the double notation for L_f which is *equal* to L_{df}. Now (3.15) allows us to compute the commutator of L_{ω_1} and L_{ω_2}:

$$\begin{aligned}[L_{\omega_1},L_{\omega_2}](f) - L_{d\Lambda(\omega_1,\omega_2)}(f) &= \Lambda(\omega_1,d\Lambda(\omega_2,df)) \\ &\quad + \Lambda(\omega_2,d\Lambda(df,\omega_1)) + \Lambda(df,d\Lambda(\omega_1,\omega_2)) \\ &= -d\omega_1(\lambda_{\omega_2},\lambda_{df}) - d\omega_2(\lambda_{df},\lambda_{\omega_1}) . \end{aligned} \tag{4.8}$$

If we restrict ω_1 and ω_2 to *closed* covariant vector fields, then

$$[L_{\omega_1},L_{\omega_2}] = L_{d\Lambda(\omega_1,\omega_2)} . \tag{4.9}$$

Equation (4.9) calls for a better notation. We introduce as bracket between covariant vector fields ω_1, and ω_2 by

$$\{\omega_1,\omega_2\} = d\Lambda(\omega_1,\omega_2) . \tag{4.10}$$

This new bracket does in general not satisfy a Jacobi identity. It does so however for *closed* vector fields. The closed vector fields therefore form again a Lie ring and we have

Theorem 4. The Lie ring of the closed covariant vector fields ω is mapped by $\omega \rightarrow L_\omega$ on a subring of $\mathfrak{L}$:

$$[L_{\omega_1},L_{\omega_2}] = L_{\{\omega_1,\omega_2\}} ; \quad d\omega_1 = 0 , \quad d\omega_2 = 0 . \tag{4.11}$$

[3] For example, E. Goursat, *Leçons sur le problème de Pfaff* (Herman & Cie, Paris, 1928).

Next we want to discuss one parameter groups in a canonical manifold.

Definition 1. A one-parameter group Φ_t is canonical if it satisfies

$$\{f \circ \Phi_t, g \circ \Phi_t\} = \{f,g\} \circ \Phi_t . \tag{4.12}$$

Such a group therefore induces a one parametric family of automorphisms in $(V,\mathcal{E},\Lambda)$. Φ_t is uniquely determined by its infinitesimal generator

$$L(f) = (d/dt)(f \circ \Phi_t)|_{t=0} . \tag{4.13}$$

Differentiation of (4.12) leads to

$$\{L(f),g\} + \{f,L(g)\} = L(\{f,g\}) ,$$

or, according to (3.9),

$$\Delta L(f,g) = 0 . \tag{4.14}$$

This however leads, due to (3.14), to

$$d\omega_L = 0 \tag{4.15}$$

and therefore

$$L = L_\omega , \qquad d\omega = 0 . \tag{4.16}$$

Theorem 5. The infinitesimal generators of one parameter group of canonical transformations are of the form L_ω, where ω is *closed.*

The function $f(t) = f \circ \Phi_t$ satisfies a differential equation

$$df/dt = \Lambda(\omega,df) , \qquad d\omega = 0 . \tag{4.17}$$

If ω is not only closed but exact ($\omega = dh$), then (4.17) can also be written in the form

$$df/dt = \{h,f\} . \tag{4.18}$$

An equation of the form (4.18) is *globally Hamiltonian.* An equation of the form (4.17) is *locally Hamiltonian* if ω is not exact.

Not any Hamiltonian equation leads to a one-parameter group of canonical transformations. Alternatively formulated: not every L_ω with $d\omega = 0$ is an infinitesimal generator of a one-parameter group of canonical transformations. In the Appendix we give a simple example of this sort and a completely trivial example of a locally Hamiltonian group of canonical transformations.

Here is a simple sufficient criterium in order that L_h is an infinitesimal generator of a group of canonical transformations

Criterium. L_h generates a group of canonical transformations by (4.18), if for any $E, \{p | h(p) \leqq E\}$ is compact.

The proof, which depends on the fact, that h itself is an integral of (4.17) can easily be adapted from Nomizu, Ref. 2.

Finally we discuss *canonical mappings*. Let $(V,\mathcal{E},\Lambda)$ and $(V',\mathcal{E}',\Lambda')$ be two canonical manifolds.

Definition 2. A mapping Φ from V into V' is canonical, if it satisfies

$$\{f'\circ\Phi,g'\circ\Phi\} = \{f',g'\}\circ\Phi\,. \tag{4.19}$$

(4.19) can also be written as

$$\Lambda(d(f'\circ\Phi),d(g'\circ\Phi))(p) = \Lambda'(df',dg')(p')\,. \tag{4.20}$$

Now we know from Sec. 1(D) that Φ induces a linear mapping of the tangent space T_p of V into the tangent space $T'_{p'}$ of V' [Eq. (1.4)]:

$$\lambda'_{p'} = \Phi'(p)\lambda_p\,. \tag{4.21}$$

$\Phi'(p)$ therefore maps the dual space $T'^*_{p'}$ into T^*_p by

$$\omega_p = \omega'_{p'}\Phi'(p) \tag{4.22}$$

and specifically

$$df_p = df'_{p'}\Phi'(p)\,. \tag{4.23}$$

Equation (4.20) therefore can be rewritten as

$$\Lambda(df'_{p'}\Phi'(p),dg'_{p'}\Phi'(p)) = \Lambda'(df'_{p'},dg'_{p'}) \tag{4.24}$$

and this generalizes (dropping p and p') to

$$\Lambda(\omega'\Phi',\eta'\Phi') = \Lambda'(\omega',\eta')\,. \tag{4.25}$$

We claim now that the kernel of the mapping Φ' contains only $\omega'=0$. In fact if $\omega'\Phi'=0$ then (4.25) vanishes for all η'. Since Λ' is nonsingular, this is only possible if $\omega'=0$. From this it follows that the mapping (4.21) is *onto* $T'_{p'}$. $\Phi'(p)$ has therefore always maximal rank, namely rank $\Phi' = \text{Dim } V'$.

Theorem 6. If Φ is a canonical mapping of $(V,\mathcal{E},\Lambda)$ into $(V',\mathcal{E}',\Lambda')$ then rank $\Phi' = \dim V'$. An open set in V is mapped onto an open set in V'.

Let us now discuss a canonical mapping of V onto itself. If such a mapping has in inverse it follows from theorem 6 that this inverse is differentiable and therefore a cononical transformation.

Theorem 7. A one-to-one canonical mapping of $(V,\mathcal{E},\Lambda)$ onto itself is a canonical transformation.

5. ORIENTATION OF A CANONICAL (OR SYMPLECTIC) MANIFOLD. INTEGRATION. LIOUVILLE'S THEOREM

We need the following well-known result:

Lemma 4. Let $\xi^{ik} = -\xi^{ki}$; $i,k = 1, 2,\cdots, 2f$ be $f(2f-1)$ variables and ξ be the matrix $||\xi^{ik}||$. The determinant Det ξ of ξ is then the square of a polynomial Pf ξ in the variable ξ^{ik}:

$$\text{Det } \xi = [\text{Pf } \xi]^2\,. \tag{5.1}$$

The *Pfaffian* Pf f is uniquely determined by the normalization Pf $\epsilon = 1$, where $\epsilon^{12} = \epsilon^{34} = \cdots = \epsilon^{2f-1,2f} = 1$ and all other $\epsilon^{kl} = 0$ for $k \leqq l$. If in addition $\xi = M\xi' M^T$, where M is a $2f\times 2f$ matrix, then

$$\text{Pf } \xi = \text{Det } M \text{ Pf } \xi'\,. \tag{5.2}$$

For a proof see, e.g., E. Artin, *Geometric Algebra* (Interscience Publishers, Inc., New York, 1957), p. 141.

Let $(V,\mathcal{E},\Lambda)$ be a manifold with symplectic structure, $N(p)$ be a cubic neighborhood to the local coordinates x^i. By

$$\eta^{ik}(q) = \{x^i,x^u\} = -\eta^{ki}(q) \tag{5.3}$$

we define a nonsingular C^∞ skew symmetric matrix η in $N(p)$. The Pfaffian of η is denoted by γ:

$$\gamma(q) = \text{Pf } \eta\,. \tag{5.4}$$

$\gamma(q)$ in $N(p)$ is also C^∞. The same is true for

$$\alpha(q) = \gamma(q)/|\gamma(q)|\,. \tag{5.5}$$

According to (2) a change of local coordinates will multiply $\gamma(q)$ by the Jacobian J and $\alpha(q)$ by

$$\alpha'(q) = (J/|J|)\alpha(q)\,. \tag{5.6}$$

$\alpha(q)$ depends therefore only on the orientation of the local coordinate system.

Definition 3. A local coordinate system has the orientation $\alpha(q)$.

The existence of a continuous $\alpha(p)$, which satisfies $\alpha^2 = 1$ and transforms according to (5.6) is equivalent to the orientability of V.[4]

We have therefore

Theorem 8. Every symplectic manifold (and therefore every canonical manifold) is orientable.

From now on we restrict the discussion to canonical manifolds. Let Φ be a canonical mapping of V into V. Let $p' = \Phi(p)$. As we know from theorem 6 Φ' is nowhere singular. Φ has therefore locally an inverse. Let $N'(p')$ be a cubic neighborhood of p' in which Φ^{-1} exists and let $N(p) = \Phi^{-1}(N'(p'))$. Finally let $y^i = f^i$ be local coordinates in $N'(p')$. Then $x^i = f^i\circ\Phi$ are local coordinates in $N(p)$. However $\{y^i,y^k\}(q') = \{x^i,x^k\}(q)$. The local coordinate systems x and y have therefore the same orientation and we have

Theorem 9. A canonical mapping of V into V never changes the orientation.

[4] G. de. Rham, *Variétés differentiables* (Hermann & Cie, Paris, 1955), Sec. 5.

In order to prepare the way to the Liouville theorem we have to introduce an invariant integration. This is easily done for functions with support in a cubic neighborhood $N(p)$. Let supp $f \subset N(p)$ and let again x^i be the corresponding local coordinates. Let finally $\hat{\eta}$ be the inverse of the matrix $\eta = ||\{x^i,x^k\}||$. According to the properties of the Pfaffian mentioned in lemma 4 one easily convinces oneself that

$$I(f) = \int \hat{f}(x) \mid \mathrm{Pf}\hat{\eta}(x) \mid dx^1 \ldots dx^{2f} \qquad (5.7)$$

is independent of the special choice of local coordinates. The generalization of (7) to more general functions (as far as it is feasible) follows the standard procedure, using a suitable decomposition of 1 (Ref. 4, Sec. 5). Thus an invariant integral can be defined, e.g., for all functions with compact support. Liouville's theorem is now almost trivial.

Theorem 10 (*Liouville*). Let Φ be a canonical transformation and g integrable. Then

$$I(g) = I(g \circ \Phi) \,. \qquad (5.8)$$

Proof. It is sufficient to prove (5.8) for functions g for which supp $g \subset N'(p')$ where $p' = \Phi(p)$. Let $y^i = f^i$ be local coordinates in $N'(p')$, then $x^i = f^i \circ \Phi$ are local coordinates in $N(p) = \Phi^{-1}(N'(p'))$. We have furthermore $\hat{g}(x) = (g \circ \Phi)\hat{}(x)$ and finally $\{y^i,y^k\}(q') = \{x^i,x^k\}(q)$. Thus (5.8) reduces to an identity.

ACKNOWLEDGMENT

Some years ago, I obtained from Professor Georges Mackey his course on Mathematical Foundations of Quantum Mechanics, which has since appeared as a book [G. W. Mackey, *Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, Inc., New York, 1963)]. The first chapter of this book deals with classical mechanics and has some contact with the material presented here. Despite a difference in attitude, it has strongly influenced my views. I would like to express my gratitude to Professor Mackey for making his work available to me.

APPENDIX (EXAMPLES)

A. Hamilton Equation Without One-Parameter Group

$$V = R^2 = \{(p,q)\}$$

Canonical structure:

$$\{f,g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\,.$$

Take $h = \frac{1}{4}p^2 - 4q^3$ with the corresponding canonical equations $\dot{q} = \{h,q\} = \frac{1}{2}p$, $\dot{p} = \{h,p\} = 12q^2$. The solutions are given in terms of Weierstrass' $\wp$ function:

$$q = \wp(t_0 - t;0, -E) = \varphi(t;q_0,p_0)$$
$$p = -2\wp'(t_0 - t;0, -E) = \psi(t;q_0,p_0)$$

where

$$E = \tfrac{1}{4}p_0^2 - 4q_0^3\,, \quad t_0 = \int_{q_0}^{\infty} \frac{ds}{(4s^3 + E)^{\frac{1}{2}}}\,.$$

$\wp(u;g_2,g_3)$ is defined by

$$u = \int_{\wp(u)}^{\infty} \frac{ds}{(4s^3 - g_2 s - g_3)^{\frac{1}{2}}}$$

and becomes singular for $u = 0$. The formulas for q and p are therefore only meaningful for $t < t_0$ and can not define a one-parameter group of canonical transformation. The physical reason for this is clear: the repulsive potential $-4q^3$ is so strong that a particle vanishes at plus infinity within a finite time which becomes the shorter, the farther to the right the particle already was at the beginning.

B. Compact Canonical Manifold

$$V = S^1 \times S^1 \quad \text{(product of two circles)}$$
$$\mathcal{E}: f(p+1,q) = f(p,q+1) = f(p,q)$$

C^∞ in p and q.

$$\Lambda : \{f,g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\,.$$

It is not clear to me whether compact canonical manifolds play any role in mechanics. Treated in statistical mechanics they would lead to negative absolute temperatures.

C. Locally but not Globally Hamiltonian Group

Take example 2:

$$\frac{df}{dt} = \alpha\frac{\partial f}{\partial q} - \beta\frac{\partial f}{\partial p} = \Lambda(\omega,df)$$

generates the group

$$(f \circ \Phi_t)(p,q) = f(p - \beta t, q + \alpha t)$$

which belongs to $\omega = \alpha dp + \beta dq$. ω is however not exact on V since $\alpha p + \beta q \notin \mathcal{E}$. Similar cases clearly occur in mechanics.