On the Reduction of Dynamical Systems by Means of Parametrized Invariant Relations
Author(s): E. R. van Kampen and Aurel Wintner
Source: *Transactions of the American Mathematical Society*, Vol. 44, No. 2 (Sep., 1938), pp. 168-195
Published by: American Mathematical Society
Stable URL: http://www.jstor.org/stable/1989969
Accessed: 25/06/2014 09:15

ON THE REDUCTION OF DYNAMICAL SYSTEMS BY MEANS OF PARAMETRIZED INVARIANT RELATIONS*

BY

E. R. van KAMPEN AND AUREL WINTNER

Introduction. The classical reduction theory of canonical systems with n degrees of freedom assumes that there is known a function group of first integrals.† The reduction of the degree of freedom is then carried out by using the Hamilton-Jacobi theory or an equivalent approach. In the present paper a more general problem will be considered, since it will not be assumed that the known functions, or hypersurfaces in the phase-space, are represented by first integrals. In fact, the first integrals will be replaced by invariant relations,‡ so that, in particular, the known hypersurfaces need not form a continuous family, but may be isolated.

It should be mentioned that the generalization of the reduction problem to the case where the given first integrals are replaced by invariant relations is not an artificial problem but one which arises quite naturally in the simplest applications. If, for instance, one wants to reduce the degree of freedom of the problem of three bodies by means of the classical first integrals in an explicit and symmetrical manner, one is compelled to replace these first integrals by certain of their combinations which form a complicated system of invariant relations and not a system of first integrals.§

The treatment of the general case of invariant relations will differ from the usual treatment of the case of first integrals in that all considerations will be based on a parametrization of the system of invariant relations, this parametrization being symmetrical with respect to the n coordinates and n impulses. Needless to say, the usual treatment of first integrals, based on the Hamilton-Jacobi theory, cannot be applied in the general case under consideration.

In the particular case where the system of invariant relations is a system of first integrals (§10), the treatment of the general case goes over into a

reduction theory which seems to have essential formal advantages over the usual treatments of this particular case.

In the limiting case where the system of invariant relations is empty, the considerations go over into a treatment of canonical transformations which has been recently developed* and which suggested the approach of the present paper to the reduction problem.

§11, which is independent of the rest of the paper, attempts an application of the method of parametrization to the case where the given relations are not invariant systems but constraints. However, the usual rule for the introduction of constraints will be verified to be identical with the rule suggested by the method of parametrization only in the case where the constraints are holonomic.†

Notation. A matrix with k rows and l columns will be called a (k, l)-matrix, so that the product CD of a (k, l)-matrix C and an (l, m)-matrix D is a (k, m)-matrix. If C is a (k, l)-matrix, C' will denote the (l, k)-matrix which is the transposed of C. A $(k, 1)$-matrix will be termed a k-vector, so that, if C and D are k-vectors, their scalar product may be represented as either of the two matrix products $C'D$, $D'C$. The (k, l)-matrix in which all kl elements are 0 will be denoted by $0^k{}_l$; and $0^k{}_1$, by 0^k. The letters $k, r, \cdots$ will denote scalar functions. Total differentiation with respect to the time t will be marked by a dot, while t as a subscript will refer to partial differentiation. If, for instance, c is a scalar function $c(Z, t)$ of a k-vector Z and of t, and if Z is a function $Z(t)$ of t, then

$$\dot{c} = c_t + \dot{Z}' \operatorname{grad}_Z c,$$

where the subscript of grad refers to the space in which one carries out the partial differentiations. The functions in question will be assumed to possess, in the regions under consideration, continuous derivatives of the μth order, where either $\mu = 1$ or $\mu = 2$ or $\mu = 3$; it will be always clear from the connection which of these three values is the actual value of μ.

1. **Parametrizations of invariant systems.** For a fixed integer $n \geq 1$, let G denote the $(2n, 2n)$-matrix

(1)
$$G = \begin{pmatrix} 0^n{}_n & -E_n \\ E_n & 0^n{}_n \end{pmatrix},$$

where E_n denotes the n-rowed unit matrix, so that

(2)
$$G = -G' = -G^{-1}.$$

---

* Wintner [7], van Kampen and Wintner [2].

† As to the holonomic case, cf. Levi-Civita [5].

The canonical system with n degrees of freedom and with the Hamiltonian function

$$h = h(q_1, \cdots, q_n, p_1, \cdots, p_n, t),$$

that is, the system

$$\dot{q}_\lambda = \frac{\partial h}{\partial p_\lambda}, \qquad \dot{p}_\lambda = -\frac{\partial h}{\partial q_\lambda}, \qquad \lambda = 1, \cdots, n,$$

may clearly be written in the form

(3) $$G\dot{X} = \operatorname{grad}_X h,$$

where X denotes the $2n$-vector

(4) $$X = (x_1, \cdots, x_{2n})$$

defined by

(4 bis) $$X = \begin{pmatrix} Q \\ P \end{pmatrix}, \qquad Q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad P = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \equiv \begin{pmatrix} x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix}$$

so that

(5) $$h = h(X, t).$$

For a fixed m, which is independent of (X, t) and is such that

(6) $$m \leqq 2n,$$

let K be a $(2n-m)$-vector defined as a function $K(X, t)$, in the $(2n+1)$-dimensional (X, t)-region under consideration, in such a way that

(7) $$2n - m = \operatorname{rank} \text{ of } \operatorname{grad}_X K,$$

and that the condition

(8) $$K \equiv K(X, t) = 0^{2n-m}$$

determines a non-empty set $\Lambda(t)$ in the $2n$-dimensional X-space for every fixed t. Since $\operatorname{grad}_X K$ is a $(2n-m, 2n)$-matrix, the assumption (7) is to the effect that $\Lambda(t)$ is a locally m-dimensional manifold for every fixed t. Hence one can choose a $2n$-vector

(9) $$F = F(Y, t),$$

depending on t and on an m-vector

(10) $$Y = (y_1, \cdots, y_m)$$

in such a way that

$$(11) \qquad\qquad X = F(Y, t)$$

is a local parametrization of the manifold $\Lambda(t)$, and that

$$(12) \qquad\qquad m = \text{rank of } J,$$

where J denotes the Jacobian $(2n, m)$-matrix

$$(13) \qquad\qquad J = J(Y, t) = \text{grad}_Y F.$$

If $Y = Y(t)$ is an arbitrary curve in the space of the m parameters (10), the corresponding curve

$$(14) \qquad\qquad X(t) = F(Y(t), t)$$

in the space of the $2n$ phase coordinates (4) is such that the relation

$$(15) \qquad\qquad \dot{X} = F_t + J\dot{Y}$$

is an identity in t.

The system (8) of $2n - m$ relations is called an *invariant system of* (3) if every solution $X(t)$ of (3) which is on the manifold $\Lambda(t)$ for a single t is on $\Lambda(t)$ for every t, where the manifold $\Lambda(t)$ is supposed to be a non-empty set. Any parametrization (11) of an invariant system (8) of (3) will also be called an invariant system of (3). An invariant system (8) of (3) does not determine the function (9) uniquely, since (12) is the only restriction as to the choice of the parameters (10). In what follows, the parametrization (11) of (8) will be considered as fixed.

It is clear from (12) that if a given solution $X(t)$ of (3) satisfies the invariant system (8), then there exists in the space of the parameter vector Y exactly one path $Y = Y(t)$ for which (14) is valid. According to (15) and (3), this unique $Y(t)$ is such that the relation

$$(16) \qquad\qquad GJ\dot{Y} = -GF_t + \text{grad}_X h$$

is, in virtue of (11), an identity in t. Now there exists for every point of the manifold $\Lambda(t)$ exactly one solution $X(t)$ of (3) passing through this point. Hence (16) and (12) imply that there belongs to the parametrization (11) of the invariant system (8) of (3) a unique m-vector

$$(17) \qquad\qquad L = L(Y, t)$$

which is defined in the $(m+1)$-dimensional (Y, t)-region in such a way that the relation

$$(18) \qquad\qquad GJL = -GF_t + \text{grad}_X h$$

is, in virtue of (11), an identity in Y and t. Since both $Y(t)$ and $L(Y, t)$ are uniquely determined, comparison of (16) with (18) shows that

$$(19) \qquad\qquad \dot{Y} = L(Y, t).$$

Accordingly, $X(t)$ is a solution of (3) satisfying the invariant system (8) of (3) if and only if the system (19) of m ordinary differential equations has a solution $Y(t)$ by means of which $X(t)$ is representable in the form (14). It is clear, from the proof, that the two differential equations (19), (16) are equivalent, although (19) has, and (16) need not have, the normal form of ordinary differential equations.

2. **Some formal consequences of the properties of** G. On defining an m-vector R by

$$(20.1) \qquad\qquad R = R(Y, t) = \tfrac{1}{2} J'GF$$

and a scalar r by

$$(20.2) \qquad\qquad r = r(Y, t) = \tfrac{1}{2} F'_t GF,$$

one will be able to show that the matrix function $J'GJ$ of (Y, t) may be represented in the (Y, t)-region under consideration as the alternating derivative (curl)

$$(21.1) \qquad\qquad J'GJ = (\mathrm{grad}_Y\, R) - (\mathrm{grad}_Y\, R)',$$

while the vector function JGF_t of (Y, t) appears in the form

$$(21.2) \qquad\qquad J'GF_t = R_t - \mathrm{grad}_Y\, r.$$

In fact, (21.1) and (21.2) are the identities

$$2J'GJ = \mathrm{grad}_Y\, (J'GF) - (\mathrm{grad}_Y\, (J'GF))'$$

and

$$2J'GF_t = (J'GF)_t - (\mathrm{grad}_Y\, (F'_t GF))'$$

which follow from (2) and (13), since the second derivatives cancel.

Since the Jacobian (13) is a $(2n, m)$-matrix, it is clear from (1) that $J'GJ$ is a skew-symmetric (m, m)-matrix, so that

$$(22) \qquad\qquad 2l = \text{rank of } J'GJ$$

for some integer l which satisfies the inequality

$$(23) \qquad\qquad 2l \leqq m$$

and will be supposed to be independent of (Y, t) in the $(m+1)$-dimensional

(Y, t)-region under consideration. It may be mentioned that (6) and (25) can be replaced by the sharper statement that

(23 bis) $$0 \leq 2l \leq m \leq n + l \leq 2n.$$

While the definition (23) of the integer l assumes, in view of (13), the choice of a parametrization (11) of the invariant system (8), it will be shown that l is uniquely determined by the invariant system (8) alone. This is implied by the relation

(24) $$2m - \text{rank of } J'GJ = 2n - \text{rank of } NGN',$$

where N denotes the $(2n - m, 2n)$-matrix

(25) $$N = N(X, t) = \text{grad}_X K,$$

so that NGN' is, in view of (1), a skew-symmetric $(2n - m, 2n - m)$-matrix. In fact, since (11) is a parametrization of (8), $K(F(Y, t), t) = 0^{2n-m}$ is an identity in (Y, t). On differentiating this identity with respect to the m components of Y, one sees from (26) and (13) that

(26.1) $$NJ = 0^{2n-m}$$

is an identity in (Y, t) in virtue of (11). Furthermore, it is seen from (12) that there exists an $(m, 2n)$-matrix $T = T(Y, t)$ such that

(26.2) $$TJ = E_m,$$

where E_m is the m-rowed unit matrix. Now it is clear from (7), (25), (26.1) and (26.2) that the $(2n, 2n)$-matrix

$$\begin{pmatrix} N \\ T \end{pmatrix}$$

has the rank $2n$; that is, it has a non-vanishing determinant; and its reciprocal matrix is, by (26.1) and (26.2), of the form

$$\begin{pmatrix} N \\ T \end{pmatrix}^{-1} = (WJ),$$

where $W = W(Y, t)$ is some $(2n, 2n - m)$-matrix. Hence (2) implies that the two $(2n, 2n)$-matrices

$$H_1 = \begin{pmatrix} N \\ T \end{pmatrix} G \begin{pmatrix} N \\ T \end{pmatrix}' = \begin{pmatrix} NGN' & NGT' \\ TGN' & TGT' \end{pmatrix}$$

and

$$H_2 = (WJ)'G(WJ) = \begin{pmatrix} W'GW & W'GJ \\ J'GW & J'GJ \end{pmatrix}$$

are reciprocal matrices. Now a well known theorem states that, since H_1 and H_2 are reciprocal matrices, the difference of order and rank in NGN' is the same as in $J'GJ$. This proves (24). A direct proof of (24) follows by multiplication of H_1 and the matrix

$$\begin{pmatrix} E_{2n-m} & W'GJ \\ 0^m{}_{2n-m} & J'GJ \end{pmatrix}.$$

The Lagrange bracket

$$(27.1) \qquad [u, v] = \sum_{\lambda=1}^{n} \frac{\partial(a_{n+\lambda}, a_\lambda)}{\partial(u, v)} = - [u, v]$$

of a $2n$-vector

$$(a_1, \cdots, a_{2n}) = A = A(u, v),$$

which depends on two scalar parameters u, v, may be written in the form

$$(27.2) \qquad [u, v] = A_u' GA_v,$$

so that the well known identity

$$(27.3) \qquad [u, v]_w + [v, w]_u + [w, u]_v = 0,$$

which holds for an arbitrary $2n$-vector $A = A(u, v, w)$ depending on three scalar parameters, may be written as

$$(27.4) \qquad (A_u' GA_v)_w + (A_v' GA_w)_u + (A_w' GA_u)_v = 0.$$

Similarly, the Poisson parenthesis

$$(28.1) \qquad (f_1, f_2) = \sum_{\lambda=1}^{n} \frac{\partial(f_1, f_2)}{\partial(x_{n+\lambda}, x_\lambda)} = - (f_2, f_1)$$

of two scalars

$$f_1 = f_1(X, t), \qquad f_2 = f_2(X, t)$$

may be written as

$$(28.2) \qquad (f_1, f_2) = (\text{grad}_X f_1)'G(\text{grad}_X f_2).$$

It may be mentioned that the matrix function $J'GJ$ of (Y, t) is independent of t if and only if the vector function $J'GF_t$ of (Y, t) is the Y-gradient of a scalar function of (Y, t). In fact, $(J'GJ)_t$ is the alternating derivative (curl) of $J'GF_t$ with respect to Y. In order to verify this, it is sufficient to replace in (27.4) the variables A, u, v, w by F, t and two components of Y, respectively.

3. **Separating parametrizations of invariant systems.** It will be convenient to split the m-vector (10), with the use of the integer (22) satisfying (23), into the $2l$-vector.

(29.1) $$Y^0 = (y_1, \cdots, y_{2l})$$

and the $(m-2l)$-vector

(29.2) $$Y^* = (y_{2l+1}, \cdots, y_m),$$

so that

(29.3) $$(y_1, \cdots, y_m) = Y = \begin{pmatrix} Y^0 \\ Y^* \end{pmatrix}.$$

The limiting case

(29 bis) $$2l = m, \quad \text{that is,} \quad Y = Y^0,$$

will not be excluded.

The parametrization (11) of (8) will be called a *separating parametrization* if there exists a $(2l, 2l)$-matrix $S^0 = S^0(Y, t)$ by means of which the (m, m)-matrix (21.1) may be represented in the form

(30) $$J'GJ = \begin{pmatrix} S^0 & 0^{2l}{}_{m-2l} \\ 0^{m-2l}{}_{2l} & 0^{m-2l}{}_{m-2l} \end{pmatrix}.$$

It is clear from (22) and from the skew-symmetry of (21.1) that

(31) $$\det S^0 \neq 0$$

and

(32) $$S^0 = -(S^0)',$$

whenever S^0 exists.

It will be shown that if $S^0 = S^0(Y, t)$ exists, that is, if (11) is a separating parametrization of the invariant system (8) of (3), then there exist a $2l$-vector

(33.1) $$R^0 = R^0(Y^0, t)$$

and a scalar

(33.2) $$h^0 = h^0(Y^0, t),$$

such that the $(2l, 2l)$-matrix S^0 occurring in (30) may be represented as the alternating derivative

(34.1) $$S^0(Y^0, t) \equiv S^0 = (\text{grad}_{Y^0} R^0) - (\text{grad}_{Y^0} R^0)',$$

while the m-vector (21.2) appears in the form

$$(34.2) \qquad J'GF_t = \text{grad}_Y\, h + \begin{pmatrix} R^0{}_t - \text{grad}_{Y^0}\, h^0 \\ 0^{m-2l} \end{pmatrix},$$

where

$$(35) \qquad h = h(Y;t) \equiv h(F(Y,t),t),$$

cf. (5), (11). This implies, in particular, that in case of a separating parametrization the functions

$$J'GJ, \ J'GF_t - \text{grad}_Y\, h$$

of t and of the m-vector (29.3) are both independent of the $(m-2l)$-vector (29.2). In the proof use will be made of the identity

$$(36) \qquad J'\,\text{grad}_X\, h = \text{grad}_Y\, h,$$

which is obvious from (11), (13), and (35).

First, it is clear that the vector function R is determined by the conditions (21.1), (21.2) only up to an additive term of the form

$$(37) \qquad \text{grad}_Y\, b,$$

where $b = b(Y,t)$ is an arbitrary scalar function. If (37) is added to R, then the term b_t has to be added to r in order to satisfy (21.2). On the other hand, on placing

$$(38) \qquad R = \begin{pmatrix} R^0 \\ R^* \end{pmatrix},$$

so that $R^0 = R^0(Y,t)$ and $R^* = R^*(Y,t)$ are the $2l$-vector and the $(m-2l)$-vector formed by the first $2l$ and the last $m-2l$ components of the m-vector R, one can see from (29.3) that (21.1) implies the pair of identities

$$(39.1) \qquad (\text{grad}_{Y^0}\, R^0) - (\text{grad}_{Y^0}\, R^*)' = 0^{2l}{}_{m-2l},$$

and

$$(39.2) \qquad (\text{grad}_{Y^*}\, R^*) - (\text{grad}_{Y^*}\, R^*)' = 0^{m-2l}{}_{m-2l},$$

if the condition (30) for a separating parametrization is satisfied. Since (39.2) is precisely the integrability condition for the existence of a scalar

$$(40) \qquad b = b(Y,t) = b(Y^0, Y^*, t)$$

such that

$$(41) \qquad R^* = \text{grad}_{Y^*}\, b$$

is an identity in $(Y, t) = (Y^0, Y^*, t)$, and since (38) is undetermined up to an additive term of the form (37), it is clear that, on modifying the functions R and r in a suitable manner, one can assume without loss of generality that the function (40) is identically zero. Then $R^* = 0^{m-2l}$ by (41); hence

$$(42) \qquad\qquad R = \begin{pmatrix} R^0 \\ 0^{m-2l} \end{pmatrix}$$

by (38), and (39.1) goes over into

$$(42 \text{ bis}) \qquad\qquad \text{grad}_Y \cdot R^0 = 0^{2l}{}_{m-2l}.$$

Now it is seen from (42 bis) and (29.3) that the function $R^0 = R^0(Y, t)$ is of the form (33.1).

In order to prove the existence of the function (33.2), notice first that on multiplying (18) by $-J'$ and using (36), one obtains

$$(43) \qquad\qquad -J'GJL = J'GF_t - \text{grad}_Y h.$$

Since substitution of (42) into (21.2) gives

$$(44) \qquad J'GF_t - \text{grad}_Y h = \begin{pmatrix} R^0{}_t - \text{grad}_{Y^0} (r + h) \\ 0^{m-2l} - \text{grad}_{Y^*} (r + h) \end{pmatrix},$$

it is seen from (30) and (43) that

$$(45) \qquad\qquad 0^{m-2l} = \text{grad}_{Y^*} (r + h).$$

Now (45) means that the function $r+h$ of $(Y, t) = (Y^0, Y^*, t)$ is independent of Y^*; that is, that the scalar function h^0 defined by

$$(46) \qquad\qquad h^0 = r + h$$

is of the form (33.2). Finally, substitution of (45) and (46) into (44) gives (34.2), while (34.1) follows from (30) if one substitutes (33.1) and (42) into (21.1).

The theorem thus proved and the last remark in §2 imply that in the separating case characterized by (30) the function (34.1) of (Y^0, t) is independent of t if and only if the function (33.1) is also, or, more precisely, that

$$(47) \qquad\qquad S^0{}_t = 0^{2l}{}_{2l} \quad \text{if and only if} \quad R^0{}_t \equiv 0^{2l}$$

holds for a *suitable* choice of the function (33.1).

4. **The separated differential equations belonging to a separating parametrization of an invariant system.** The name "separating parametrization" will now be justified by showing that if the parametrization (11) of an invari-

ant system (8) of (3) is a separating parametrization, then the first $2l$ of the m differential equations represented by (19) appear in the form

$$(48) \qquad \dot{Y}^0 = L^0(Y^0, t),$$

that is, as a system of $2l$ differential equations which is separated from the $m - 2l$ differential equations determining the components (29.2) of the m-vector (29.3) as functions of t. The latter differential equations are, in view of (19) and (29.3), of the form

$$(49) \qquad \dot{Y}^* = L^*(Y^*, Y^0, t).$$

Thus if there is known a solution $Y^0(t)$ of (48), the corresponding functions $Y^*(t)$ are the solutions of

$$(50) \qquad \dot{Y}^* = L^*(Y^*, t),$$

where

$$(50 \text{ bis}) \qquad L^*(Y^*, t) = L^*(Y^*, Y^0(t), t)$$

(cf. (49)) is a known function of t and of the $(m - 2l)$-vector (29.2), so that (19) splits into the two separated systems (48), (50). Finally, the m-vector $Y(t)$, determined by $Y^0(t)$ and $Y^*(t)$ in virtue of (29.3), gives the corresponding solution (14) of (3). The system (48) of order (22) will be referred to as the separated system belonging to the separating parametrization (11) of (8).

In order to prove the possibility of splitting (19) into (48) and (50), notice first that (19) is, as pointed out at the end of §1, equivalent to the system (16). Now, if (16) is multiplied by J', the first $2l$ of the resulting m differential equations appear, in view of (30), (34.2), (36), and (29.3), as

$$(51) \qquad S^0\dot{Y}^0 = -R^0{}_t + \operatorname{grad}_{Y^0} h^0.$$

This completes the proof, since (33.1), (33.2), (34.1), and (31) imply that (51) can be written in the form (48). Needless to say, (48) is in the limiting case (29 bis) identical with the original system (19). The result just proved contains, however, information for this limiting case also.

In fact, whether (29 bis) is or is not satisfied, the result proved above may be described by saying that if the parametrization (11) of an invariant system (8) of (3) is a separating parametrization, then the $2l$-vector (29.2) is determined by a system of $2l$ differential equations which form a *non-degenerate Pfaffian dynamical system* with l degrees of freedom.[†] In order to see this, it

---

[†] The term *Pfaffian dynamical system* is used in the sense of Birkhoff. Such a system is called non-degenerate if it can be solved with respect to the time derivatives, that is, if the skew-symmetric matrix, represented by the curl of the vector potential involved, does not vanish in the domain of the phase space which is under consideration.

is sufficient to notice that on the one hand (34.1) and (51) are equivalent to

$$(52) \qquad \{(\mathrm{grad}_{Y^0} R^0) - (\mathrm{grad}_{Y^0} R^0)'\}\dot{Y}^0 = -R^0{}_t + \mathrm{grad}_{Y^0} h^0,$$

and that on the other hand (52) clearly is the system of Euler-Lagrange equations which belongs to the calculus of variations problem

$$(52 \text{ bis}) \qquad \delta \int \{(\dot{Y}^0)' R^0(Y^0, t) + h^0(Y^0, t)\} dt = 0,$$

where the ends are fixed and the problem is non-degenerate in view of (31) and (34.1).

5. **Existence of separating parametrizations of an invariant system.** The considerations of §3 and §4 assumed that there is *given* a separating parametrization (11) of the invariant system (8) of (3). Actually, one can always *choose* the parametrization of an invariant system as a separating parametrization. This existence theorem can be inferred from a classical theorem in the theory of Pfaffians as follows.

First, (21.1), (21.2), (36), and (16) hold for arbitrary parametrizations of (8). Let (11) be a given parametrization. Since from (16) and (36)

$$(53) \qquad J'GJ\dot{Y} + J'GF_t - \mathrm{grad}_Y h = 0^m,$$

it is seen from (21.1) and (21.2) that

$$(54) \qquad \{(\mathrm{grad}_Y R) - (\mathrm{grad}_Y R)'\}\dot{Y} + R_t - \mathrm{grad}_Y (r + h) = 0^m,$$

where R, r, and h are the functions (20.1), (20.2), and (35). If one denotes by ω the $(m+1)$-ary Pfaffian

$$(55) \qquad \omega = R' dY + (r + h) dt$$

(cf. (10)), it is clear that (53) is the system of Euler-Lagrange equations belonging to

$$(56) \qquad \delta \int \omega \equiv \delta \int \{[R(Y, t)]' dY + [r(Y, t) + h(Y, t)] dt\} = 0,$$

where the ends are fixed. Notice that (56), in contrast with (52 bis), need not be non-degenerate, since the determinant of the skew-symmetric (m, m)-matrix (21.1) vanishes identically unless (29 bis) is satisfied.

Let G^0 denote the $(2l, 2l)$-matrix

$$(57) \qquad G^0 = \begin{pmatrix} 0^l{}_l & -E_l \\ E_l & 0^l{}_l \end{pmatrix} = -(G^0) = -(G^0)^{-1},$$

obtained from (1), (2) by replacing n by l, where l is defined by (22) and satis-

fies (23 bis). It is well known† that on subjecting the m-dimensional Y-space to a suitable non-singular local point transformation for every fixed t, and omitting then an additive complete differential, one can transform the Pfaffian (55) into the $(2l+1)$-ary normal form

$$(58) \qquad\qquad -\tfrac{1}{2}Y^{0\prime}G^0 dY^0 + h^0 dt,$$

it being understood that in (58) the notation (29.3) is applied to the *new* m-vector Y and that

$$(59) \qquad\qquad h^0 = h^0(Y^0, Y^*, t)$$

is some scalar function. Since a variation problem (56) is covariant under point transformations, and since a complete differential can always be omitted beneath the signs $\delta\!\!\int$, where the ends are fixed, it is clear from (59) and from the normal form (58) of (55) that (56) appears in terms of the new m-vector Y in the form

$$(60) \qquad\qquad \delta\int \left\{ -\tfrac{1}{2}Y^{0\prime}G^0 dY^0 + h^0(Y^0, Y^*, t)dt \right\} = 0.$$

Now, (57) being a constant matrix, it is seen from (29.3) that the Euler-Lagrange equations belonging to (60) are

$$(61) \qquad \begin{pmatrix} G^0 & 0^{2l}{}_{m-2l} \\ 0^{m-2l}{}_{2l} & 0^{m-2l}{}_{m-2l} \end{pmatrix} \begin{pmatrix} \dot{Y}^0 \\ \dot{Y}^* \end{pmatrix} - \begin{pmatrix} \mathrm{grad}_{Y^0}\, h^0 \\ \mathrm{grad}_{Y^*}\, h^0 \end{pmatrix} = 0^m.$$

On substituting (29.3) into (57) and comparing the result with (61), one obtains

$$(62) \qquad\qquad J'GJ = \begin{pmatrix} G^0 & 0^{2l}{}_{m-2l} \\ 0^{m-2l}{}_{2l} & 0^{m-2l}{}_{m-2l} \end{pmatrix}.$$

Consequently, the condition (30), which characterizes the separating parametrizations, is satisfied, S^0 being the matrix (57). This completes the proof of the existence theorem stated at the beginning of this section.

Let it be mentioned for application in §6 that the function (59) occurring in (61) can be chosen, in view of the results of §3, as a function of the form (33.2). In other words, one can write (61) in the form

$$(63) \qquad\qquad G^0\dot{Y}^0 = \mathrm{grad}_{Y^0}\, h^0$$

in view of $\mathrm{grad}_{Y^*}\, h^0 \equiv 0^{m-2l}$.

---

† The normal form usually given in textbooks is not (58) but another Pfaffian. However, the latter differs from the Pfaffian (58) only in a complete differential; cf., e.g., van Kampen and Wintner [2], p. 862.

6. **Canonical parametrizations of an invariant system.** A given parametrization (11) of a system (8) of m invariant relations of a canonical system (3) with n degrees of freedom will be called a *canonical parametrization* if it is a separating parametrization (§3), and the corresponding separated system of $2l$ differential equations, that is, the system (48), is a canonical system with l degrees of freedom, so that (52) is of the particular form (63). Thus the result found at the end of §5 may be expressed by saying that the invariant system (8) of (3) always admits a parametrization which is a canonical parametrization.

The construction of a canonical parametrization of a given invariant system depends on the construction of a point transformation which transforms the Pfaffian (55) into its normal form (58). And the construction of such a point transformation is known to require the solution of a complete system of partial differential equations, a solution which, in general, cannot be obtained by mere quadratures. Hence, if one does not want to refer to the existence theorem of complete systems, which depends on successive approximations or on equivalent processes, one has to assume that the point transformation in question is a priori known. This is the situation in many applications where the geometrical or dynamical connection suggests the "correct" choice of the parameters (10). Actually, it is easy to see from the considerations of §5 that the point transformation which transforms the Pfaffian (55) into its normal form (58) can be chosen such as to depend only on the function (8) and not on the function (5). In other words, (8) admits a parametrization (11) which is a *fixed* canonical parametrization for *all* those canonical systems (3), with n degrees of freedom, for which the *fixed* system (8) of m relations is an invariant system.

It will be assumed in what follows that both (8) and (3) are fixed. The considerations of §5 clearly imply that if a given separating parametrization (11) of (8) is such that (30) is of the particular form (62), that is, if

$$(64) \qquad\qquad S^0 = G^0$$

holds for the constant $(2l, 2l)$-matrix (57) and for the matrix (34.1) occurring in the definition (30) of a separating parametrization, then the given separating parametrization is a canonical parametrization.

If a canonical parametrization (11) of an invariant system (8) of (3) is such that, for some constant m-vector C, the Y_0-gradient of the Hamiltonian function (33.2) of the separated system (63) is the 0^m-vector for every t, then

$$Y^0(t) \equiv C$$

clearly is a solution of (63). For this solution of equilibrium, (50) goes over, in view of (50), into

(65)
$$\dot{Y}^* = L^*(Y^*, C, t),$$

so that

(66)
$$Y(t) = \begin{pmatrix} Y^0(t) \\ Y^*(t) \end{pmatrix} \equiv \begin{pmatrix} C \\ Y^*(t) \end{pmatrix}$$

is, for every solution $Y^*(t)$ of (65), a solution of (19). The solutions $X(t)$ of (3) obtained from (66.1) by means of (14) are, in the main, the stationary solutions of Routh and Levi-Civita.†

7. **Completely canonical parametrizations.** A canonical parametrization (11) of an invariant system (8) of (3) will be said to be *completely canonical* if the Hamiltonian function (33.2) of the separated canonical system (63) may be obtained by direct substitution of (11) and (29.3) into the Hamiltonian function (5) of the original canonical system (3), so that

(67)
$$h^0 = h^0(Y^0, t) \equiv h(F(Y^0, Y^*, t), t).$$

This implies, of course, that $h(F(Y^0, Y^*, t), t)$ is a function of Y^0 and t alone, that is, that

(67 bis)
$$\mathrm{grad}_{Y^*} h(F(Y^0, Y^*, t), t) \equiv 0^{m-2l}.$$

If a canonical parametrization (11) of (8) does not contain t, that is, if $F_t \equiv 0$, then the parametrization is completely canonical. In order to prove this, notice first that if $F_t \equiv 0^{2n}$, then, as seen from §3, the functions R and r of Y and t are independent of t. Hence it is sufficient to prove that r is independent of Y. In fact, r is then a constant and can, therefore, be omitted in (46). Now $\mathrm{grad}_Y r \equiv 0^m$ follows from (21.2) in view of $R_t \equiv 0^m$ and $F_t \equiv 0^{2n}$.

Remark. On comparing the fact thus proved with the considerations of §5, it is seen that every invariant system (8) which does not contain t explicitly admits a completely canonical parametrization (the Hamiltonian function (5) need not be independent of t).

As an application of completely canonical parametrizations, a generalization of the well known case of ignorable coordinates will now be considered.

If a coordinate $x_\lambda = q_\lambda$ (cf. (4), (4 bis)) does not occur in the Hamiltonian function (5) of (3), the corresponding impulse $x_{n+\lambda} = p_\lambda$ clearly is independent of t for *every* solution $X(t)$ of (3). Thus the case of an ignorable coordinate is the case of first integrals (cf. §10) of the form $x_\lambda = \mathrm{const.}$ and is implied by the more general case where $x_\lambda = q_\lambda$ does occur in (5) but $x_{n+\lambda} = p_\lambda$ has a constant

† Cf. Levi-Civita [4].

value $c_{n+\lambda}$ for *some* solutions $X(t)$ of (3), so that $x_{n+\lambda}=c_{n+\lambda}$ is not a first integral but merely an invariant relation. This case of coordinates which are ignorable in a generalized sense is again a particular case of the one where a system (8) of $2n-m$ invariant relations is given as consisting on the one hand of $n-l$ relations of the form

$$(68.1) \qquad\qquad x_{n+\lambda} = c_{n+\lambda}, \qquad\qquad \lambda = l + 1, \cdots, n,$$

and on the other hand of a system of $m-n+l$ relations of the type

$$(68.2) \qquad x_\mu = f_\mu(x_1, \cdots, x_{m-l}; x_{n+1}, \cdots, x_{2n}), \qquad \mu = m - l + 1, \cdots, n,$$

where l denotes some integer satisfying (23 bis). The last assumption implies that

$$(68 \text{ bis}) \qquad\qquad 0 \leqq l \leqq n \quad \text{and} \quad m - l \leqq n$$

and that all the x's occurring in (68.1) are impulses p in view of (4 bis). The case (68.1), (68.2) of (8) is of importance in certain applications† and can be parametrized in a completely canonical way as follows:

First, a parametrization (11) of the invariant system (8) represented by the m relations (68.1) and (68.2) may be obtained by choosing as the m components y of the parameter vector (10) those m components x_i of the $2n$-vector (4) for which either $1 \leqq i \leqq m-l$ or $n \leqq i \leqq n+l$. If one numerates the m components of this Y in a suitable manner, the parametrization (11) of the invariant system (68.1), (68.2) appears in the form (11), if one adjoins the relations

$$(68.3) \quad x_i = y_i, x_{n+i} = y_{l+i}, x_j = y_{l+j}, i = 1, \cdots, l; j = l + 1, \cdots, m - l,$$

to (68.1), (68.2). Hence it is easily verified from (1) and (13) that the integer l occurring in (68.1), (68.2), (68.3) is the same as the one defined by (22), and that the matrix $J'GJ$ is, in view of (57), precisely the matrix (62). Since (62) is satisfied, the parametrization is canonical; that it is completely canonical follows from the fact that t does not occur in (68.1), (68.2), (68.3), that is, in (11).

Accordingly, the reduction of (3) to the separated canonical system (63) is, in the case of an invariant system of the form (68.1), (68.2), identical with the reduction of (3) in the classical case of first integrals represented by ignorable coordinates. In this particular case, the function (50 bis) is, in view of (68.3), independent of Y^*; hence $Y^* = Y^*(t)$ follows from (50) by quadrature, while the system (50) determining Y^* cannot be solved always by quadratures in the more general case (68.1), (68.2), (68.3).

---

† Cf. van Kampen and Wintner [3], pp. 155–156 and pp. 164–166.

8. **Canonical parametrization of holonomic invariant systems.** In this section the number l will be introduced as a given integer for which

(69) $$0 \leqq l \leqq n,$$

but it will turn out that l is identical with the number defined by (22), while the number m defined by (7) will become

(70) $$m = n + l.$$

The invariant system (8) of (3) will be called *holonomic* if it is independent of the n components of the $2n$-vector (4) which are represented in (4 bis) by the n-vector P; that is, if (8) is of the form

(71) $$U(Q, t) = 0^{n-l},$$

where

(72) $$U = U(Q, t)$$

is an $(n-l)$-vector depending on the n-vector Q, defined by (4 bis) and possibly on t. It is clear from (4 bis) that in this holonomic case (7) does over into

(73) $$n - l = \text{rank of } I,$$

where I denotes the Jacobian $(n-l, m)$-matrix

(74) $$I = I(Q, t) = \text{grad}_Q U,$$

m being the number (70). According to (70), one can write the parameter vector (10) in the form

(75) $$Y = \begin{pmatrix} \overline{Q} \\ \overline{P} \\ P* \end{pmatrix},$$

where $\overline{Q}$ and $\overline{P}$ are l-vectors and $P*$ is an $(n-l)$-vector. Furthermore, it is seen from (73) that (71) admits by means of an n-vector

(76) $$A = A(\overline{Q}, t)$$

a parametrization of the form

(77) $$Q = A(\overline{Q}, t),$$

where

(78) $$l = \text{rank of } \text{grad}_{\overline{Q}} A.$$

This clearly implies the existence of (l, n)-matrices

$$(79) \qquad\qquad\qquad H = H(\overline{Q}, t)$$

for which

$$(80) \qquad\qquad\qquad H \operatorname{grad}_{\overline{Q}} A = E_l$$

(where E_l is the unit matrix) is an identity.

Now it will be shown that if an invariant system (8) of (3) is holonomic, then it admits a parametrization (11) which is canonical and of the form

$$(81) \qquad\qquad \binom{Q}{P} \equiv X = F(Y, t) = \binom{A}{H'\overline{P} + I'P^*},$$

where the m-vector Y is written in the form (75), the n-vector A is arbitrarily chosen such that (77) gives a parametrization of (71) satisfying (78), and H is any matrix satisfying (80); finally the matrix I defined by (74) is thought of as expressed by means of (77) as a function of $(\overline{Q}, t)$.

In order to prove this, notice first that substitution of (77) into (71) gives an identity in $(\overline{Q}, t)$. On differentiating this identity with respect to the components of the l-vector $\overline{Q}$, one obtains

$$(82) \qquad\qquad\qquad I \operatorname{grad}_{\overline{Q}} A = 0^{n-l}{}_l$$

in view of (74). Now (80), (82), and (73) clearly imply that the (n, n)-matrix

$$(83) \qquad\qquad\qquad \binom{H}{I}$$

has the rank n; that is, it has a non-vanishing determinant. Hence it is clear from (75) that, since (77) is a parametrization of the restriction (71) for Q, (81) is a parametrization of the invariant system (8) represented by (71) and (4 bis). It remains to be shown that this parametrization of the holonomic invariant system is canonical.

First, the $(2n, n+l)$-matrix J defined by (13) is, according to (81) and (75),

$$J = \begin{pmatrix} \operatorname{grad}_{\overline{Q}} A & 0^n{}_l & 0^n{}_{n-l} \\ \operatorname{grad}_{\overline{Q}} (H'\overline{P} + I'P^*) & H' & I' \end{pmatrix}.$$

Hence it is seen from (1), (57) and (80), (82) that the matrix product $J'GJ$ satisfies the criterion (62) of a canonical parametrization if and only if the (l, l)-matrix

$$(\operatorname{grad}_{\overline{Q}} (H'\overline{P} + I'P^*))' \operatorname{grad}_{\overline{Q}} A - (\operatorname{grad}_{\overline{Q}} A)' \operatorname{grad}_{\overline{Q}} (H'\overline{P} + I'P^*)$$

is identically zero, that is, if and only if the (l, l)-matrix

$$D = (\text{grad}_{\bar{Q}} A)' \, \text{grad}_{\bar{Q}} (H'\overline{P} + I'P^*)$$

is symmetric. Now if one brings the factor $(\text{grad}_{\bar{Q}} A)'$ of D beneath the grad sign of the second factor of D, one clearly obtains

$$D = \text{grad}_{\bar{Q}} \left\{ (\text{grad}_{\bar{Q}} A)'(H'\overline{P} + I'P^*) \right\} + \text{correction term,}$$

where the correction term, being an iterated gradient, is a symmetric matrix. Hence it is sufficient to show that

$$\text{grad}_{\bar{Q}} \left\{ (\text{grad}_{\bar{Q}} A)'(H'\overline{P} + I'P^*) \right\}$$

is a symmetric matrix. Now this matrix is

$$\text{grad}_{\bar{Q}} \left\{ (H \, \text{grad}_{\bar{Q}} A)'\overline{P} + (I \, \text{grad}_{\bar{Q}} A)P^* \right\} = \text{grad}_{\bar{Q}} \, \overline{P}$$

in view of (80) and (82); hence it is the zero matrix, since $\overline{P}$ is independent of $\overline{Q}$ (cf. (75)).

9. **Canonical parametrizations of semi-holonomic invariant systems.** The two types of invariant systems treated below under (i) and (ii) are important in some applications.† Each of these two invariant systems consists of two subsystems one of which is independent of the impulses while the other is linear in the impulses. One of the two subsystems, when considered without the other, need not be an invariant system. The number l will be introduced, in both cases (i), (ii), as a given integer satisfying (69); but it turns out that l is identical with the number l defined by (22), while (70) is replaced by

(84) $$m = 2l.$$

The difference between the "semi-holonomic" cases (i), (ii) and the holonomic case of §8 depends, in the main, on the fact that in the semi-holonomic cases the $(n-1)$-vector P^* is missing in (75), so that the m-vector (10) is, in accordance with (84), of the form

(85) $$Y = \begin{pmatrix} \overline{Q} \\ \overline{P} \end{pmatrix},$$

where Q and P are l-vectors.

   Case (i). Let

(86.1) $$U = U(Q, t),$$

(86.2) $$V = V(Q, t)$$

_____

† Cf. van Kampen and Wintner [3], pp. 154–155, and the applications given in [3].

be a pair of functions such that U is an $(n-l)$-vector, V is an l-vector and, if I and H denote the Jacobian $(n-l, n)$- and (l, n)-matrices

$$(87.1) \qquad\qquad I = I(Q, t) = \mathrm{grad}_Q\, U,$$

$$(87.2) \qquad\qquad H = H(Q, t) = \mathrm{grad}_Q\, V,$$

the determinant of the (n, n)-matrix (83) is distinct from zero, which clearly implies that

$$(88.1) \qquad\qquad n - l = \text{rank of } I,$$

$$(88.2) \qquad\qquad l = \text{rank of } H.$$

Hence there exist $(n-l, n)$-matrices

$$(89) \qquad\qquad M = M(Q, t)$$

such that

$$(90) \qquad\qquad H M' = 0^l{}_{n-l}$$

is an identity in (Q, t) and

$$(91) \qquad\qquad n - l = \text{rank of } M.$$

The non-vanishing of the determinant of the (n, n)-matrix (83) also implies that, if $\overline{Q}$ denotes the l-vector

$$(92) \qquad\qquad \overline{Q} = V(Q, t),$$

there exists a unique n-vector (76) such that (78) is satisfied and (77) gives a parametrization (71) in such a way that (92) is an identity in $(\overline{Q}, t)$ in virtue of (77).

Now suppose that an invariant system (8) of (3) is given by the pair of conditions

$$(93.1) \qquad\qquad U = 0^{n-l},$$

$$(93.2) \qquad\qquad M P = 0^{n-l},$$

where U, M are given functions (86.1), (89) which satisfy, for some function (86.2), the conditions described above, and Q, P are the n-vectors defined by (4 bis). It will be shown that the invariant system (8) represented by (93.1), (93.2) admits a parametrization (11) which is canonical and of the form

$$(94) \qquad\qquad \begin{pmatrix} Q \\ P \end{pmatrix} \equiv X = F(Y, t) \equiv \begin{pmatrix} A \\ H'\overline{P} \end{pmatrix},$$

where the m-vector Y is written in the form (85) by means of two l-vectors

$\overline{Q}$, $\overline{P}$, ($m = 2l$); the unique n-vector (76) defined before by means of (92) is such that (77) gives a parametrization of (93.1); the matrix H defined by (87.2) is thought of as expressed by means of (77) as function of (Q, t); and finally the inversion formula belonging to (93) is

$$(95) \qquad Y \equiv \begin{pmatrix} \overline{Q} \\ \overline{P} \end{pmatrix} = \begin{pmatrix} V \\ (\mathrm{grad}_{\overline{Q}}\, A)'\overline{P} \end{pmatrix}$$

in virtue of the invariant system.

In fact, the first row of (95) and the first row of (94) are satisfied by the definitions (92) and (77) of $\overline{Q}$ and A. Furthermore, it is seen from the definition (87.2) of H that (80) is an identity in $(\overline{Q}, t)$ in virtue of (77). Now (80), (90), and (91) imply that the (n, n)-matrix

$$(96) \qquad \begin{pmatrix} (\mathrm{grad}_{\overline{Q}}\, A)' \\ M \end{pmatrix}$$

has the rank n, that is, has a non-vanishing determinant. Hence the system

$$(97) \qquad \begin{pmatrix} (\mathrm{grad}_{\overline{Q}}\, A)' \\ M \end{pmatrix} P = \begin{pmatrix} \overline{P} \\ 0^{n-l} \end{pmatrix}$$

of n linear equations has, for any l-vector $\overline{P}$, a unique solution P; and this P is

$$(97 \text{ bis}) \qquad P = H'\overline{P}$$

in view of (80) and (90). Hence, on considering the first row of (97) as the definition of $\overline{P}$, the second row of (94) follows from (93.2). This proves that the invariant system (8) represented by (93.1), (93.2) admits the parametrization (94) and satisfies (95). In order to prove that this parametrization is a canonical parametrization, that is, that the $(2l, 2l)$-matrix $J'GJ$, where J is defined by (13) and (94), satisfies the identity (62), one merely has to repeat, with obvious modifications, the verification carried out in detail at the end of §8.

Case (ii). Suppose that an invariant system (8) of (3) is given by a pair of conditions

$$(98.1) \qquad U = 0^{n-l},$$

$$(98.2) \qquad MP = 0^{n-l},$$

where Q, P are the n-vectors defined by (4 bis), and

$$(99.1) \qquad U = U(Q, t),$$

$$(99.2) \qquad M = M(Q, t),$$

(U is an $(n-l)$-vector, M is an $(n-l, n)$-matrix), and

(100) $n - l = \text{rank } IM'$,

where I denotes the $(n-l, n)$-matrix (74). The existence of a function (86.2), occurring in Case (i), is not assumed in the present Case (ii). Assumption (100) clearly implies (74) and (91), while (73) implies the existence of n-vectors (76) which satisfy (78) and are such that (77) gives a parametrization (98.1) in terms of an l-vector $\overline{Q}$.

It will be shown that the invariant system (8) represented by (98.1), (98.2) admits, for every choice of the function (76), a parametrization (11) which is canonical and of the form (94), (85), where the l-vector $\overline{P}$ is such that

(101) $\overline{P} = (\text{grad}_{\overline{Q}} A)'P,$

while $H = H(\overline{Q}, t)$ is an (l, n)-matrix which is uniquely determined by the choice of (76) and satisfies (80) and (90).

In order to prove this, notice first that (82) is an identity in virtue of (77), the reason being the same as in §8. Now (82), (77), and (100) imply that the (n, n)-matrix (96) has the rank n, that is, it has a non-vanishing determinant. Hence there exists a unique (l, n)-matrix $H = H(\overline{Q}, t)$ which satisfies (80) and (90). Let (101) be considered as the definition of an l-vector $\overline{P}$. Since the determinant of (96) is distinct from zero, the linear system (97) has a unique solution P; and this P is given by (97 bis) in view of (80) and (90). This proves the second row of (94), while the first row of (94) is identical with (77). Finally, (98.2) and (101) are satisfied in view of (97). This proves that (98) is a parametrization (11) of the invariant system (8) represented by (98.1), (98.2). That this parametrization is a canonical parametrization, is proved in the same way as in Case (i).

10. **The case of first integrals.** The results of the preceding sections concern the general case of an invariant system (8) of (3). In what follows, it will be assumed that the relation

(102) $K(X, t) = C,$

where C is a constant $(2n-m)$-vector, is an invariant system of (3) not only for a single choice of the constant vector C but for all choices of C in some $(2n-m)$-dimensional region of the C-space. This is the case if and only if each of the $2n-m$ scalar functions which constitute the components of the vector function $K = K(X, t)$ are first integrals of (3). Since a scalar function $g = g(X, t)$ represents a first integral of (3) if and only if $\dot{g} = 0$ is an identity in virtue of (3) alone, it is sufficient to prove that if (102) is an invariant system for every C, then $\dot{K} = 0^{2n-m}$ is an identity in virtue of (3) alone. Now if (102)

is an invariant system for every C, then $\dot{K} = 0^{2n-m}$ is an identity in virtue of (3) and (102) together, and so, C being arbitrary, it is an identity in virtue of (3) alone.

In the particular case where the invariant system consists of first integrals, the integer $2l$ defined by (22) or (24) has a simple meaning, since it is connected with an integer occurring in the classical theory of reduction of canonical systems by means of first integrals. In fact, it turns out that if the $2n - m$ first integrals represented by (102), which are independent in view of (7), form a function group in the sense of Lie,* then this function group contains $n - l$, but not more, independent first integrals in involution. In view of the well known extension process of Poisson parentheses, the assumption that the $2n - m$ first integrals form a function group does not imply any loss of generality.

Since, on placing

(103) $2k = $ rank of NGN',   where $N = \mathrm{grad}_x\, K$,

one can write (24), in view of (22), in the form

(104) $n - k = m - l$,

the statement to be proved may be formulated as follows: If (102) consists of $2n - m$ first integrals of a canonical system (3) with n degrees of freedom, and if these first integrals are independent and form a function group, then

(105.1) $j = 2n - m - k$,

where k is the integer defined by (103) and j denotes the maximum number of first integrals in involution which are contained in the function group.

First, NGN' is a skew-symmetric $(2n - m,\ 2n - m)$-matrix in which the j elements contained in the first j rows and first j columns vanish, if the first j components of the $(2n - m)$-vector $K(x, t)$ are independent first integrals in involution. Hence it is clear from (103) that

(105.2) $2k \leqq 2(2n - m - j)$,   that is, $j \leqq 2n - m - k$.

In order to complete the proof of (105.1), it remains to be shown that

(105.3) $j \geqq 2n - m - k$.

Now (105.3) is implied by a known existence theorem concerning functions in involution.†

---

    * Cf., e.g., the presentation of Engel [1].
    † The presentation of the local existence theorems by Engel [1] assumes, but actually does not use at all, that the systems are analytic.

Remark. There may be mentioned an essential difference between the case (8) of invariant systems and the case (102) of first integrals (in involution). In the case of first integrals one can solve (50) by mere quadratures. In other words, the separation of (19) into (48) and (50) reduces the degree of freedom from n to l in case of first integrals. This is a well known consequence of the Hamilton-Jacobi theory (Lie). On the other hand, the solution of (50) can require operations higher than quadratures in case of invariant systems.

Application to the problem of three bodies. In order to illustrate the use of (104) and (105.1), consider the non-planar case of the problem of three bodies. Let ξ_i, η_i, ζ_i, where $i = 1, 2, 3$, be Cartesian coordinates of the mass μ_i in an inertial coordinate system, and let Ξ_i, H_i, Z_i denote the conjugated impulses $\mu_i\dot{\xi}_i$, $\mu_i\dot{\eta}_i$, $\mu_i\dot{\zeta}_i$. Then $n = 9$ and (3) admits, besides the energy integral, the nine classical first integrals represented by the functions

$$f_1 = \sum_{i=1}^{3} \{H_i\zeta_i - Z_i\eta_i\}, \quad f_2 = \sum_{i=1}^{3} \{Z_i\xi_i - \Xi_i\zeta_i\}, \quad f_3 = \sum_{i=1}^{3} \{\Xi_i\eta_i - H_i\xi_i\};$$

$$f_4 = \sum_{i=1}^{3} \Xi_i, \qquad f_5 = \sum_{i=1}^{3} H_i, \qquad f_6 = \sum_{i=1}^{3} Z_i;$$

$$f_7 = \sum_{i=1}^{3} \{\mu_i\xi_i - t\Xi_i\}, \quad f_8 = \sum_{i=1}^{3} \{\mu_i\eta_i - tH_i\}, \quad f_9 = \sum_{i=1}^{3} \{\mu_i\zeta_i - tZ_i\}.$$

Let the $(2n-m)$-vector K occurring in (105.1) be defined as the 9-vector with the components $f_1, \cdots, f_9$, so that $m = 9$. Thus (104) and (105.1) go over into

$$(106) \qquad\qquad l = k \quad \text{and} \quad j = 9 - k.$$

Now it will be easy to verify from (103) that $k = 4$, so that $l = 4$ and $j = 5$ are, respectively, the degree of freedom of the separated system (63) and the maximum number of first integrals in involution contained in the function group generated by $f_1, \cdots, f_9$.

In order to prove that $k = 4$, one has to show in view of (103) and (25.1) that the 9-rowed skew-symmetric matrix

$$\|(f_\rho, f_\sigma)\|; \qquad\qquad \rho, \sigma = 1, \cdots, 9,$$

of Poisson parentheses has the rank 8 if the values of the coordinates ξ_i, η_i, ζ_i and the impulses Ξ_i, H_i, Z_i do not have exceptional values. Now it is easily verified from the above definition of $f_1, \cdots, f_9$ that

$$\begin{array}{lll} (f_1, f_2) = f_3, & (f_2, f_4) = -f_6, & (f_3, f_4) = f_5, \\[2mm] (f_2, f_7) = -f_9, & (f_3, f_7) = f_8, & (f_4, f_7) = \sum_{i=1}^{3} \mu_i, \end{array}$$

and that these 6 identities remain valid under simultaneous cyclic permuta-
tions of the three triples $(1, 2, 3)$, $(4, 5, 6)$, $(7, 8, 9)$; and finally that those
$9^2 - 2 \cdot 18 = 45$ Poisson parentheses which occur neither among these eighteen
nor among their negatives $(f_\sigma, f_\rho) = -(f_\rho, f_\sigma)$, are identically zero. In other
words, the skew-symmetric $(9, 9)$-matrix $\|(f_\rho, f_\sigma)\|$ may be written in the form

$$(107) \qquad \|(f_\rho, f_\sigma)\| = \begin{pmatrix} F_1 & F_2 & F_3 \\ F_2 & 0 & \sum_{i=1}^{3} \mu_i E \\ F_3 & -\sum_{i=1}^{3} \mu_i E & 0 \end{pmatrix},$$

where 0 and E denote the 3-rowed zero and unit matrices, respectively, and
F_i is the three-rowed skew-symmetric matrix

$$(108) \qquad F_i = \begin{pmatrix} 0 & f_{3i} & -f_{3i-1} \\ -f_{3i} & 0 & f_{3i-2} \\ f_{3i-1} & -f_{3i-2} & 0 \end{pmatrix}, \qquad i = 1, 2, 3.$$

It is easily seen from (108) and from the above representation of $f_1, \cdots, f_9$
that, if the eighteen coordinates and impulses have values which do not lie
on an algebraical manifold of a dimension number less than eighteen, the
9-rowed skew-symmetric matrix (107) is of rank 8; that is, $k = 4$.

    If one considers the planar solutions of the problem of three bodies, one
can assume that the three ζ_i and the three Z_i are identically zero, and there
remain only the five integrals f_3, f_4, f_5, f_7, f_8. Thus $n = 6$ and $2n - m = 5$, so
that (104) and (105.1) go over into

$$(109) \qquad\qquad l = k + 1, \; j = 5 - k.$$

Since the 5-rowed skew-symmetric matrix which takes the place of the 9-
rowed matrix $\|(f_\rho, f_\sigma)\|$ is found to be of rank 4 by the above method, it fol-
lows that $k = 2$, and therefore, by (109), that $l = 3$ and $j = 3$.

    11. Forced paths. So far it has been assumed that (8) is an invariant sys-
tem of (3). In the present section it will only be assumed that (8) satisfies (7)
and determines a non-empty manifold $\Lambda(t)$; that is, that (8) is compatible
with itself, without necessarily being compatible with (3); so that (8) plays
the role of a system of $2n - m$ constraints which modify the equations of mo-
tion (3). These constraints, in view of (4 bis), are not necessarily holonomic
and are not given in the usual way, that is, not in terms of the coordinates,
velocities, and time. It can be assumed, in view of (7), that the manifold (8)
is represented by means of a parameter vector (10) in the form (11), (12).

Denote by

(110) $$Y = Y(t)$$

any curve in the Y-space, with the time t as parameter, and let the curve in the X-space,

(111) $$X = X(t),$$

which is defined by (110) and (14), be called the image curve of (110). Put, along this image path,

(112) $$Z = Z(t) = G\dot{X} - \text{grad}_X h,$$

so that the $2n$-vector $Z(t)$ is identically zero if and only if the image curve of (110) represents a solution of (3). The $(2n, m)$-matrix (13) is a function of t along the curve (110). Now if a curve (111) in the X-space is such that it is the image curve of a curve (110) for which

(113) $$J'Z = 0^m$$

is an identity in t, then (111) will be termed a *forced path* belonging to the Hamiltonian function (5) and to the constraint system (8), since it is clear that whether a path (111) is a forced path or not is independent of the choice of the parametrization (11) of (8). It is seen from the identities (15), (36) and from the definitions (112) and (113) that a path (111) is a forced path belonging to (5) and (8) if and only if the differential condition

(114) $$J'GJ\dot{Y} = -J'GF_t + \text{grad}_Y h,$$

where h is defined by (35), has a solution (110) by means of which (111) is representable in the form (14). Since the skew-symmetric (m, m)-matrix $J'GJ$, defined in the $(m+1)$-dimensional (Y, t)-region under consideration, can have a rank (22) which is less than m, the system (114) cannot always be written in the normal form (19) of ordinary differential equations. In other words, one has general existence and uniqueness theorems for the initial problem of (114) only when accidentally $m = 2l$.

In what follows, only Hamiltonian functions (5) and constraints (8) will be considered for which the initial problem of (114) has not only an existence but also a uniqueness theorem in the region under consideration (to this end, the assumption $m = 2l$ is sufficient but not necessary). Then the dynamical meaning of forced paths is indicated by the following remark: *If one assumes that the $2n$-vector (112) is a force, then the orthogonality condition (113) defining a forced path is nothing but d'Alembert's principle; which expresses the*

fact that the forces Z, which modify the equations (3) of free motion in view of the constraint (8), do not work.

It is hard to verify the general validity of this interpretation. The interpretation is certainly valid in case (8) is not a constraint but an invariant system, since in this case (16) is valid and (114) follows from (36). Furthermore, the remark is valid in case the constraints are holonomic, since in this case (113) will be seen to be identical with the differential equations usually derived from d'Alembert's principle by means of Lagrangian multipliers. Whether or not the assumption (113) is dynamically correct in case of non-holonomic constraints, is a difficult, if not meaningless, question. In fact, non-holonomic constraints are usually assigned in terms of velocities and not of impulses, while X in (8) does contain impulses in view of (4 bis). Now, in case P actually occurs in (8), the passage from the impulse vector P to the velocity vector Q cannot be carried out before solving (113).

It remains to be verified that in the case of holonomic constraints which can contain t explicitly, the requirement (103) of a forced path leads to the same paths as d'Alembert's principle; that is, in the holonomic case the forced paths are identical with the paths described under the constraints which are required by the usual rules of dynamics. First, if (8) depends only on Q and t and not on P, the equations of motion belonging to the Hamiltonian function (3) and to the system (8) of $2n - m$ constraints are

$$\dot{Q} = \mathrm{grad}_P \ h, \qquad - \dot{P} = \mathrm{grad}_Q \ h + (\mathrm{grad}_Q \ K)'\lambda,$$

where λ is a $(2n-m)$-vector representing Lagrangian multipliers. This may be written, in view of (4 bis) and (1), in the form

$$G\dot{X} = \mathrm{grad}_X \ h + (\mathrm{grad}_X \ K)'\binom{0^m}{\lambda},$$

since $\mathrm{grad}_P \ K \equiv 0^n{}_{2n-m}$. On multiplying by J' and using the definition (112) of Z, one obtains

$$J'Z = J'(\mathrm{grad}_X \ K)'\binom{0^m}{\lambda}.$$

This relation is identical with the equation (113) of forced paths, since

$$(\mathrm{grad}_X \ K)J = 0\,{}^m{}_{2n-m}$$

is, according to (13), an identity in (Y, t) in virtue of (11).

Needless to say, the consideration of forced paths in case of holonomic constraints gives for d'Alembert's principle a formulation which is invariant under canonical transformations of the $2n$-dimensional phase-space.

REFERENCES

1. F. Engel, *Die Liesche Theorie der partiellen Differentialgleichungen erster Ordnung*, edited by K. Faber, Leipzig, 1932.

2. E. R. van Kampen and A. Wintner, *On the canonical transformations of Hamiltonian systems*, American Journal of Mathematics, vol. 58 (1936), pp. 851–863.

3. E. R. van Kampen and A. Wintner, *On a symmetrical canonical reduction of the problem of three bodies*, American Journal of Mathematics, vol. 59 (1937), pp. 153–166.

4. T. Levi-Civita, *Sur la recherche des solutions particulières des systèmes différentiels et sur les mouvements stationaires*, Prace Matematyczno-Fizyczne, vol. 17 (1906), pp. 1–40.

5. T. Levi-Civita, *Sulla introduzione di vincoli olonomi nelle equazioni dinamiche di Hamilton* Atti del Reale Istituto Veneto, (8), vol. 18_1 (1916), pp. 387–395.

6. A. Wintner, *On the linear conservative dynamical systems*, Annali di Matematica, (4), vol. 13 (1934), pp. 105–112.

THE JOHNS HOPKINS UNIVERSITY,
    BALTIMORE, MD.