

# CONTRIBUTIONS TO THE TRANSFORMATION THEORY OF DYNAMICS\*

BY

DANIEL C. LEWIS, JR.†

The applications of transformation theory to dynamics are familiar through the writings of Poincaré, Levi-Civita, Hadamard, Birkhoff and others. The last named mathematician has given an extensive treatment of the so-called conservative transformations which are of particular use in the study of dynamical systems of two degrees of freedom.‡ Our present purpose is to initiate a similar treatment for those transformations in spaces of higher dimensions which are particularly important from the dynamical point of view and may be regarded as appropriate generalizations of the conservative surface transformations. These are the so-called Pfaffian transformations as defined in §1. In this paper we restrict attention to properties which are essentially characteristic of the Pfaffian transformations and not properties which the Pfaffian transformations have in common with other transformations.§ The most important results of the present paper were discovered independently by G. D. Birkhoff.|| The proofs are published now for the first time.

**1. Definition of a Pfaffian transformation and some elementary theorems.** Consider a region  $R$  of  $n$  dimensional space in which are defined  $n$  analytic functions  $X_1(x), \dots, X_n(x)$  of the  $n$  variables  $x_1, \dots, x_n$ . For the sake of brevity we write

$$a_{ij} = \partial X_i / \partial x_j - \partial X_j / \partial x_i, \quad i, j = 1, 2, \dots, n,$$

and we assume that the skew symmetric determinant  $|a_{ij}|$  is of rank  $2k$  at every point of  $R$ , where  $k$  is a positive integer not greater than  $n/2$ .

Consider also an analytic transformation  $T$ , and let us denote by  $\bar{x}_1, \dots, \bar{x}_n$  [or more briefly by  $(\bar{x})$ ] the point into which the point  $(x)$  is carried by  $T$ . It is assumed that  $T$  is defined when  $(x)$  is in  $R$ .

---

\* Presented to the Society, April 19, 1935 and September 8, 1939; received by the editors March 31, 1939.

† These results were obtained for the most part while the writer was a National Research Fellow.

‡ G. D. Birkhoff [1]. The numbers in brackets refer to the list of references at end of paper.

§ For a treatment of some of these latter properties cf. Lewis [1, 2].

|| Birkhoff [2, p. 144]. My own results were presented to the American Mathematical Society April 19, 1935 (cf. Lewis [3]) several months before Professor Birkhoff's paper was off the press, but more than a year after his results were obtained. I first learned of his results at the meeting after I had presented my own.

DEFINITION. *The transformation  $T$  is said to be Pfaffian with respect to the linear differential form*

$$(1.1) \quad \sum_{i=1}^n X_i(x) dx_i,$$

if  $\sum_{i=1}^n [X_i(\bar{x}) d\bar{x}_i - X_i(x) dx_i]$ , thought of as a differential form in the  $n$  independent variables  $x_1, \dots, x_n$ , is an exact differential, at least whenever  $(\bar{x})$  as well as  $(x)$  belong to  $R$ .

If two linear differential forms differ from each other by an exact differential, then a transformation which is Pfaffian with respect to one of them is also Pfaffian with respect to the other. *Linear differential forms differing from each other by exact differentials shall therefore be said to be equivalent to each other.\** It should be further noted that equivalent differential forms give rise to identical matrices  $(a_{ij})$ .

If

$$\sum_{i=1}^n [X_i(\bar{x}) d\bar{x}_i - X_i(x) dx_i] = \sum_{i=1}^n \left[ \sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} - X_i(x) \right] dx_i$$

is to be an exact differential, it is necessary and sufficient that

$$\frac{\partial}{\partial x_h} \left\{ \sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} - X_i(x) \right\} \equiv \frac{\partial}{\partial x_i} \left\{ \sum_{j=1}^n X_j(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_h} - X_h(x) \right\},$$

$i, h = 1, \dots, n.$

Hence we obtain the result that *a necessary and sufficient condition that  $T$  be Pfaffian with respect to (1.1) is that*

$$(1.2) \quad \sum_{j,l=1}^n a_{jl}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_h} \equiv a_{ih}(x), \quad i, h = 1, \dots, n.$$

Another necessary and sufficient condition that  $T$  be Pfaffian with respect to (1.1) is that  $\int \sum_{i=1}^n X_i dx_i$  be a relatively invariant integral of  $T$ . That is, if  $C$  is an arbitrary closed regular curve in  $R$  which is carried by  $T$  into a curve  $\bar{C}$  also in  $R$ , then the above line integral extended over  $C$  is equal to the integral extended over  $\bar{C}$  in the corresponding sense. †

That *there exist infinitely many Pfaffian transformations corresponding to an arbitrary linear differential form (1.1) may be shown in the following way:*

\* This is not the usual definition of equivalence for two differential forms. Cf. Weber [1, p. 128].

† There are in general numerous other integral invariants which can be written in compact form in the notation of Grassman's "exterior calculus." Cf. Goursat [1, pp. 211-212, 229-235]. They are derived from the relatively invariant line integral exactly as for differential equations.

Consider a system of analytic differential equations of the type

$$(1.3) \quad \frac{dx_i}{dt} = F_i(x), \quad i = 1, \dots, n.$$

The right-hand members, for simplicity, may be assumed not to involve  $t$  explicitly. Let  $x_i = f_i(t, \bar{x})$  be the solution, which for  $t=0$  takes on the initial values  $x_i = \bar{x}_i$ . These equations, for every fixed value of  $t$ , may be regarded as defining a transformation from the point  $(\bar{x})$  to the point  $(x)$ . It is known that a necessary and sufficient condition that (1.3) admit  $\int \sum_{i=1}^n X_i dx_i$  as a relatively invariant line integral is that  $\sum_{i,j} a_{ij} F_j dx_i$  be an exact differential.\* It may be easily proved that the  $F$ 's can always be chosen in an infinite number of ways so that this condition is fulfilled. The rest of the proof is left to the reader.

2. **The partial reduction of the form (1.1).** It is known from the theory of linear differential forms that, if  $2k < n$ , we may introduce a change of variables, valid in the neighborhood of an assigned point, in such a way that the form (1.1), or one equivalent to it, may be written as a form in a smaller number of variables.† By repeated application of this result, we may without loss of generality assume that (1.1) is a form in just  $2k$  variables, at least if we restrict attention to the vicinity of an invariant point of  $T$ .

Let us now suppose that the variables  $x_1, \dots, x_n$  are such that the differential form (1.1) involves only  $x_1, \dots, x_{2k}$ . In other words  $X_\alpha \equiv 0$ , ( $\alpha = 2k + 1, \dots, n$ ), and  $\partial X_i / \partial x_\alpha \equiv 0$ , ( $i = 1, \dots, n; \alpha = 2k + 1, \dots, n$ ), and hence  $a_{i\alpha} = -a_{\alpha i} = 0$ . It follows that *the transformation  $T$  (assumed to be non-singular) may be written in the form*

$$(2.1) \quad \begin{aligned} \bar{x}_l &= \bar{x}_l(x_1, \dots, x_{2k}), & l &= 1, \dots, 2k, \\ \bar{x}_\alpha &= \bar{x}_\alpha(x_1, \dots, x_{2k}, x_{2k+1}, \dots, x_n), & \alpha &= 2k + 1, \dots, n. \end{aligned}$$

*In other words the first  $2k$  of the equations defining  $T$  are independent of  $x_{2k+1}, \dots, x_n$ . The variables  $x_1, \dots, x_{2k}$  are thus said to form a "separated system."*

To prove the italicized statement, we note from (1.2) that under the present hypotheses

$$(2.2) \quad \sum_{j,l=1}^{2k} a_{jl}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_h} = a_{ih}(x), \quad i, h = 1, 2, \dots, 2k,$$

$$(2.3) \quad \sum_{j,l=1}^{2k} a_{jl}(\bar{x}) \frac{\partial \bar{x}_j}{\partial x_i} \frac{\partial \bar{x}_l}{\partial x_\alpha} = 0, \quad \alpha = 2k + 1, \dots, n.$$

\* Goursat [1, p. 219].

† Weber [1, pp. 216-217].

Here the  $4k^2$  scalar equations (2.2) may be thought of as a single matrix equation, the matrix on the right being the  $2k$ -rowed skew symmetric matrix  $(a_{ih})$  whose rank by hypothesis is  $2k$ . Hence the determinant, the element in whose  $i$ th row and  $l$ th column is  $\sum_{j=1}^{2k} a_{jl}(\bar{x})\partial\bar{x}_j/\partial x_i$ , cannot be zero. It follows from (2.3) that  $\partial\bar{x}_l/\partial x_\alpha = 0$ ,  $(l=1, \dots, 2k)$ .

If  $2k=n$ , we shall say that the Pfaffian transformation is *nondegenerate*. The significance of the theorem just proved is that, in studying the characteristic properties of Pfaffian transformations, we may for the most part confine attention to the nondegenerate case.

Apart from the fact that the form (1.1) has not yet been considered in its canonical form, the result obtained above is closely analogous to a theorem of Lie and Koenig on differential equations.\*

**3. Invariant manifolds.** Let  $T$  be a nondegenerate Pfaffian transformation, represented by the equations  $\bar{x}_i = \bar{x}_i(x)$ ,  $(i=1, \dots, n=2k)$ . Let it be assumed that  $T$  admits an  $m$ -dimensional analytic invariant manifold  $M$ , given by the equations  $x_i = f_i(u_1, \dots, u_m)$ . This means that there exists a nonsingular transformation  $S$  in  $m$ -dimensional space of the type  $\bar{u}_l = \bar{u}_l(u)$ ,  $(l=1, 2, \dots, m)$ , such that

$$f_i(\bar{u}, \dots, \bar{u}_m) \equiv \bar{x}_i[f(u_1, \dots, u_m)], \quad i = 1, \dots, 2k,$$

the identity being with respect to  $u_1, \dots, u_m$ . It is obvious that  $S$  is also a Pfaffian transformation, with respect to the linear differential form

$$\sum_{l=1}^m \left( \sum_{i=1}^{2k} X_i[f(u_1, \dots, u_m)] \frac{\partial f_i}{\partial u_l} \right) du_l \equiv \sum_{l=1}^m U_l du_l,$$

at least, if this form in the  $u$ 's is not an exact differential. Let  $b_{ij} = (\partial U_i / \partial u_j) - (\partial U_j / \partial u_i)$ . We find from an elementary calculation that  $b_{jk} = \sum a_{il} (\partial f_i / \partial u_j) (\partial f_l / \partial u_k)$ . In other words, if we let  $A$  represent the matrix  $(a_{il})$ ,  $B$  the matrix  $(b_{jk})$  and  $J$  the Jacobian matrix  $(\partial f_i / \partial u_j)$ , we have  $B = J'AJ$ , where  $J'$  is the transpose of  $J$ . The rank of  $J$  is  $m$ ; otherwise our manifold  $M$  would not be  $m$ -dimensional. It follows that the ranks of  $J'A$  and  $AJ$  are each equal to  $m$ . Letting the rank of the skew symmetric matrix  $B$  be  $2\nu$ , we find from Frobenius' theorem† on the ranks of matrices that  $\nu \geq m - k$ . Thus, if  $m > k = n/2$ , the transformation  $S$  is sure to be Pfaffian. In any case, when  $\nu > 0$ , the parameters  $u_1, \dots, u_m$  can be "separated" just as the variables  $x_1, \dots, x_n$  were separated in (2.1). It is interesting to observe that  $B$ , and hence also  $\nu$ , depends only on  $M$  and (1.1) and does not depend

\* Whittaker [1, p. 275].

† MacDuffee [1, p. 11]. The theorem is stated for square matrices; but rectangular matrices can always be made square (without change of rank) by adding rows or columns of zeros.

upon the particular transformation  $T$ , so long as it is Pfaffian with respect to (1.1) and leaves  $M$  invariant.

The foregoing results are closely analogous to some results obtained recently by Wintner and van Kampen for differential equations.\*

**4. The characteristic multipliers.** Let us now consider a nonsingular non-degenerate Pfaffian transformation in the neighborhood of an invariant point, which we take at the origin. In such a neighborhood  $T$  may be represented by power series

$$(4.1) \quad \bar{x}_i = \sum_{j=1}^{2k} c_{ij}x_j + \text{higher terms.}$$

Denoting by  $C$  the matrix of the  $c$ 's, we shall prove that *the latent roots of  $C$  occur in reciprocal pairs. That is, the  $2k$  roots  $\lambda_1, \lambda_2, \dots, \lambda_{2k}$  may be labelled so that  $\lambda_{2i-1}\lambda_{2i} = 1, (i = 1, \dots, k)$ .*

Let  $\alpha_{ij}$  be the constant term in the power series development of  $a_{ij}(x)$ . Then equating the constant terms in the identity (1.2), we obtain  $\sum_{j,l=1}^{2k} \alpha_{jl}c_{ji}c_{lh} = \alpha_{ih}$ . If we denote by  $A$  the matrix  $(\alpha_{ij})$ , these equations may be written in the form  $C'AC = A$ , where  $C'$  is the transpose of  $C$ . It follows that  $A^{-1}C'A = C^{-1}$ , where now the inverses  $A^{-1}$  and  $C^{-1}$  exist, since both  $A$  and  $C$  are nonsingular by hypothesis. Now corresponding to a latent root  $\lambda$  of  $C$ , we have a root  $\lambda^{-1}$  of  $C^{-1}$  occurring with the same multiplicity. On the other hand the above matrix equation shows that the latent roots of  $C^{-1}$  are the same as those of  $C'$  and hence also the same as those of  $C$ . The theorem readily follows, provided that we can prove that  $\det C = +1$ . Otherwise, there would be, for example, the possibility that  $C$  might admit two simple self-reciprocal roots  $+1$  and  $-1$ .

It is well known that a skew-symmetric determinant, typified by  $\det A$ , may be represented as follows:

$$2^k k! (\det A)^{1/2} = \sum_i' \epsilon^{i_1 i_2 i_3 i_4 \dots i_{2k}} \alpha_{i_1 i_2} \alpha_{i_3 i_4} \dots \alpha_{i_{2k-1} i_{2k}},$$

where  $\epsilon^{i_1 i_2 \dots i_{2k}}$  is  $+1$  or  $-1$  according as the permutation  $(i_1, \dots, i_{2k})$  of the  $2k$  integers  $(1, 2, \dots, 2k)$  is even or odd. The summation  $\sum'$  is extended over all permutations. But since  $A = C'AC$ , we have

$$\begin{aligned} 2^k k! (\det A)^{1/2} &= \sum_{i_1 \dots i_{2k}} \sum_i' \epsilon^{i_1 i_2 \dots i_{2k}} \alpha_{j_1 j_2} \dots \alpha_{j_{2k-1} j_{2k}} c_{j_1 i_1} c_{j_2 i_2} \dots c_{j_{2k} i_{2k}} \\ &= (\det C) \sum_i' \epsilon^{i_1 i_2 \dots i_{2k}} \alpha_{j_1 j_2} \dots \alpha_{j_{2k-1} j_{2k}} = (\det C) 2^k k! (\det A)^{1/2}. \end{aligned}$$

Here we use the fact that the only terms of the sum which do not cancel

\* Wintner and van Kampen [1].

each other are the terms in which the  $j$ 's are mutually distinct. We also use the elementary definition of the value of a determinant. Since  $\det A \neq 0$ , the stated result that  $\det C = +1$  follows at once.

The main result of this section is analogous to a corresponding classical result for differential equations, proved by Poincaré for the Hamiltonian case, and by G. D. Birkhoff and A. Wintner for the more general Pfaffian case.\*

**5. The formal differential system of a Pfaffian transformation.** It is known† that, if  $S$  is an arbitrary real nonsingular analytic transformation, Pfaffian or not, which leaves the origin invariant, it is possible to find a positive integer  $N$ , such that the transformation  $S^N$ , which we shall hereafter call  $T$ , has the following properties:

I.  $T^t$  (transforming the point  $x(0)$  into  $x(t)$ ) may be represented in the neighborhood of the origin (for integral  $t$ ) by means of power series of the form

$$(5.1) \quad x_i(t) = \sum_{j=1}^n y_{ij}(t)x_j(0) + \sum_{\mu=2}^{\infty} \left[ \sum_{\alpha_1+\dots+\alpha_n=\mu} y_{i\alpha_1\dots\alpha_n}(t)x_1^{\alpha_1}(0) \cdots x_n^{\alpha_n}(0) \right],$$

$i = 1, \dots, n,$

where the coefficients  $y_{ij}, y_{i\alpha_1\dots\alpha_n}$ , are entire functions of  $t$ , real when  $t$  is real, such that the series (5.1) converges when  $t$  is an integer. They do not in general converge for all  $t$ . The  $y$ 's also are such that any polynomial in a finite number of them, which vanishes for positive integral values of  $t$ , vanishes for all values of  $t$ .

II. There exist real formal power series  $Y_i(x)$  in the variables  $x_1, \dots, x_n$  which are such that the right-hand members of (5.1) satisfy (in a formal sense) the formal differential equations

$$(5.2) \quad \frac{dx_i}{dt} = Y_i(x), \quad i = 1, \dots, n,$$

and reduce to  $x_1(0), \dots, x_n(0)$ , when  $t=0$ . It should be further noted that the series  $Y_i(x)$  do not have constant terms.

We now proceed to study the formal differential system (5.2) in the case when  $S$ , and consequently  $T$ , is a Pfaffian transformation with respect to the form (1.1). We define the formal series  $U_i(x)$  by means of the formulas

$$(5.3) \quad U_i(x) = \sum_{j=1}^n a_{ij}Y_j \equiv \sum_{j=1}^n \left( \frac{\partial X_i}{\partial x_j} - \frac{\partial X_j}{\partial x_i} \right) Y_j, \quad i = 1, \dots, n.$$

\* Cf. Wintner [2].

† Lewis [1]. The integer  $N$  may, in general, be taken as unity.

We shall prove that *there exists a formal power series*  $Q(x)$ , such that

$$(5.4) \quad U_i(x) = \frac{\partial Q}{\partial x_i}.$$

Applying the identity (1.2) to the transformation  $T^t$ , for the case when  $t$  is an integer, we obtain

$$(5.5) \quad \sum_{i,l=1}^n a_{il}[x(t)] \frac{\partial x_i(t)}{\partial x_j(0)} \frac{\partial x_l(t)}{\partial x_h(0)} \equiv a_{jh}[x(0)], \quad j, h = 1, \dots, n.$$

This identity stands for an infinite number of polynomial relations connecting the  $y$ 's which hold for integral values of  $t$ . Hence by property I, we easily see that (5.5) must hold formally for all values of  $t$ . Hence, writing  $x_i$  for  $x_i(0)$ , differentiating (5.5) formally with respect to  $t$  and then setting  $t=0$ , while remembering that  $\partial x_i(0)/\partial x_h = \delta_{ih}$  and that  $dx_i(t)/dt|_{t=0} = Y_i(x)$ , we get the formal identities

$$(5.6) \quad \sum_{r=1}^n \frac{\partial a_{jh}}{\partial x_r} Y_r + \sum_{i=1}^n a_{ih} \frac{\partial Y_i}{\partial x_j} + \sum_{l=1}^n a_{jl} \frac{\partial Y_l}{\partial x_h} \equiv 0.$$

From the definition of the  $a_{ij}$  first introduced in §1, we have  $a_{ij} = -a_{ji}$  and

$$\frac{\partial a_{js}}{\partial x_h} + \frac{\partial a_{hj}}{\partial x_s} + \frac{\partial a_{sh}}{\partial x_j} \equiv 0.$$

With the help of these relations, (5.6) can readily be put into the form

$$\frac{\partial}{\partial x_h} \left[ \sum_{s=1}^n a_{js} Y_s \right] \equiv \frac{\partial}{\partial x_j} \left[ \sum_{s=1}^n a_{hs} Y_s \right].$$

Hence referring back to (5.3), we see that we have established the identities  $\partial U_j/\partial x_h \equiv \partial U_h/\partial x_j$ , ( $j, h=1, \dots, n$ ), from which the existence of a formal series  $Q$  satisfying (5.4) is obvious.

We now note that  $Q$  is *invariant under*  $T$ . For from (5.3) and (5.4) we find that

$$\sum_{i=1}^n Y_i \frac{\partial Q}{\partial x_i} \equiv \sum_{i=1}^n Y_i U_i \equiv \sum_{i,j=1}^n Y_i a_{ij} Y_j,$$

which vanishes identically in as much as  $a_{ij} = -a_{ji}$ . But the formal relation  $\sum \partial Q/\partial x_i Y_i \equiv 0$  is known to be necessary and sufficient that  $Q$  be invariant under  $T$ .

It is clear from (5.4), (5.3), and (5.2) that\*

\* The system (5.7) is equivalent to (5.2) in the nondegenerate case  $2k=n$ .

$$(5.7) \quad \sum_{j=1}^n a_{ij} \frac{dx_j}{dt} = \frac{\partial Q}{\partial x_i}, \quad i = 1, \dots, n.$$

We now ask ourselves what becomes of (5.7) if we make a change of variables. The answer is contained in the following theorem:

**THEOREM.** *Let the arbitrary nonsingular analytic transformation  $x_i = x_i(y)$ , ( $i = 1, \dots, n$ ), carry the differential form  $\sum_{i=1}^n X_i dx_i$  into  $\sum_{i=1}^n Y_i dy_i$ . For brevity, let  $b_{ij} = (\partial Y_i / \partial y_j) - (\partial Y_j / \partial y_i)$ . Then the formal differential equations (5.7) are carried over into  $\sum_{j=1}^n b_{ij} dy_j / dt = \partial Q^* / \partial y_i$ , ( $i = 1, \dots, n$ ), where  $Q^*$  is the formal series in  $y_1, \dots, y_n$  obtained by substituting in  $Q$  the power series expressions for the  $x$ 's in terms of the  $y$ 's.*

In the sequel, when no confusion can occur, the asterisk will be omitted without further comment.

**Proof.** We note the elementary relations

$$b_{ij} = \sum_{h,l} a_{hl} \frac{\partial x_h}{\partial y_i} \frac{\partial x_l}{\partial y_j}.$$

Hence

$$\sum_{j=1}^n b_{ij} \frac{dy_j}{dt} = \sum a_{hl} \frac{\partial x_h}{\partial y_i} \frac{\partial x_l}{\partial y_j} \frac{dy_j}{dt} = \sum a_{hl} \frac{dx_l}{dt} \frac{\partial x_h}{\partial y_i},$$

and substituting from (5.7) this becomes

$$\sum_{h=1}^n \frac{\partial Q}{\partial x_h} \frac{\partial x_h}{\partial y_i} = \frac{\partial Q}{\partial y_i}.$$

This completes the proof when once the obvious formal interpretation of the above symbols is supplied.†

**6. Canonical form of a Pfaffian transformation.** It is well known‡ that it is possible to choose coordinates, valid in the neighborhood of an assigned point, say the origin, in such a way that (1.1), or a differential form equivalent to it, may be written in the canonical form

$$(6.1) \quad \sum_{i=1}^k x_{2i} dx_{2i-1}.$$

Any transformation, Pfaffian with respect to (6.1), is said to be *in canonical form*. If it is nondegenerate ( $2k = n$ ) and in canonical form, it is also called

† In case  $Q$  is convergent the invariance of the equations (5.7) is obvious from the fact that they express the conditions for the vanishing of the first variation of  $\int (\sum X_i dx_i / dt + Q) dt$ . But for  $Q$  divergent this simple proof needs a reinterpretation by no means obvious.

‡ Weber [1, pp. 216-217].



a *contact transformation*. Any change of variables, which takes (6.1) into an equivalent form, will preserve the canonical form of a Pfaffian transformation by its very definition. Such a change of variables is itself a Pfaffian transformation in canonical form and is, therefore, also a contact transformation, if  $2k = n$ . On the other hand a change of variables which preserves the canonical form of a Pfaffian transformation need not be canonical nor even Pfaffian. It may, for example, change (6.1) into a constant multiple of itself.\*

In case of a contact transformation, the formal differential equations (5.2) or (5.7) take the particularly simple Hamiltonian form

$$(6.2) \quad \frac{dx_{2i-1}}{dt} = -\frac{\partial Q}{\partial x_{2i}}, \quad \frac{dx_{2i}}{dt} = \frac{\partial Q}{\partial x_{2i-1}}, \quad i = 1, \dots, k.$$

Any change of variables which preserves the canonical form of the transformation will naturally preserve the Hamiltonian form of (5.2); but it is only when the transformation defining the change of variables is a contact transformation that the Hamiltonian of the transformed equations is the original Hamiltonian with the original variables replaced by their expressions in terms of the new variables. For other transformations further recourse must be had to the general theorem at the end of §5.

Let  $W$  be an arbitrary analytic function of  $x_{2i}$  and  $y_{2i-1}$ , ( $i = 1, \dots, k$ ). Then the equations

$$(6.3) \quad y_{2i} = \frac{\partial W}{\partial y_{2i-1}}, \quad x_{2i-1} = \frac{\partial W}{\partial x_{2i}}, \quad i = 1, \dots, k,$$

are well known to define a contact transformation, provided that it is possible to solve (6.3) for the  $y$ 's in terms of the  $x$ 's and for the  $x$ 's in terms of the  $y$ 's. This is easily seen from the point of view of our present definition from the fact that

$$dW = \sum_{i=1}^k (y_{2i} dy_{2i-1} + x_{2i-1} dx_{2i})$$

is an exact differential; and hence also

$$dW - d\left(\sum_{i=1}^k x_{2i-1} x_{2i}\right) = \sum_{i=1}^k (y_{2i} dy_{2i-1} - x_{2i} dx_{2i-1})$$

is an exact differential. The transformations of this type, which leave the

---

\* There are also transformations which may preserve the canonical character of particular contact transformations without having the general property. Thus any nonsingular transformation of coordinates preserves the canonical character of the identity transformation.

origin invariant, will play an important role in the sequel. Such transformations are known to form a group.

**7. Normal form for the linear terms of a contact transformation in the case of simple elementary divisors.** Suppose that a real contact transformation  $T$ , having the origin as an invariant point, is given by power series of the form (4.1). If the matrix  $(c_{ij})$  has simple elementary divisors, it is known that it is possible to introduce a new set of variables  $(x'_1, \dots, x'_n)$  depending linearly on the original set  $(x_1, \dots, x_n)$  in such a way that  $T$  may be written in the normal form

$$(7.1) \quad \bar{x}'_j = \lambda_j x_j + \text{higher terms}, \quad j = 1, \dots, 2k = n,$$

where  $\lambda_1, \dots, \lambda_{2k}$  are the latent roots of  $(c_{ij})$ . In accordance with the results of §4, the  $\lambda$ 's may be taken so that

$$(7.2) \quad \lambda_{2i-1} = e^{\rho_i}, \quad \lambda_{2i} = e^{-\rho_i}, \quad i = 1, \dots, k.$$

Simultaneously the differential form (6.1) will be transformed into the form

$$(7.3) \quad \sum_{i,j=1}^{2k} K_{ij} x'_i dx'_j,$$

and consequently  $T$  is no longer necessarily a contact transformation in the new variables.

The variables  $(x')$  are, moreover, in general no longer real; but they may be chosen in such a way that a real point in the original  $(x)$ -space always gives rise to a real value for  $x'_j$ , if  $\lambda_j$  is real, but to conjugate imaginary values for  $x'_r$  and  $x'_s$ , if  $\lambda_r$  and  $\lambda_s$  are conjugate imaginary. Variables of this type will be said to have the *property R*.

We now ask ourselves two questions: (I) Is it possible to make a linear change of variables in such a way that the linear terms of  $T$  are reduced to the form (7.1) while (7.3) is equivalent to a (nonzero) constant multiple of  $\sum_{i=1}^k x'_i dx'_{2i-1}$ ? In other words, can we bring it about that  $T$ , when expressed in normalizing variables, is still a contact transformation? (II) Is it possible to make the transformation satisfy the conditions of question (I) and in addition be such that the new variables have the property  $R$ ?

The answer to question (I) is yes. The answer to question (II) is, in general, no, but is yes, if there are no conjugate imaginary  $\lambda$ 's of modulus unity, or if all the  $\lambda$ 's occur in conjugate imaginary pairs of modulus unity. However, even when the answer to (II) is in the negative, there always exist complex numbers  $\alpha_1, \dots, \alpha_{2k}$  such that the  $2k$  products  $\alpha_1 x'_1, \dots, \alpha_{2k} x'_{2k}$  will have

the property  $R$ . To establish these statements the following discussion is sketched. The details are left to the reader.\*

Suppose that we are already in possession of variables  $(x')$  reducing  $T$  to the normal form (7.1), transforming (6.1) to (7.3) and having the property  $R$ . We shall also at first assume that the  $\lambda$ 's are distinct. We proceed to replace (7.3) by the simplest equivalent differential form.

Since  $K_{ii}x'_i dx'_i$  is an exact differential and since  $K_{ij}x'_i dx'_j + K_{ji}x'_j dx'_i$  is equal to  $(K_{ji} - K_{ij})x'_j dx'_i$  plus the exact differential  $K_{ij}(x'_i dx'_j + x'_j dx'_i)$  and since, furthermore, in subtracting an exact differential from (7.3) we merely replace the latter by an equivalent form, we may assume that

$$(7.4) \quad K_{ji} = 0, \quad \text{whenever } j \leq i; \quad j, i = 1, 2, \dots, 2k.$$

Now the transformation given by (7.1) is Pfaffian with respect to (7.3). Hence, equating the constant terms in the identity corresponding to (1.2), we get

$$(K_{hi} - K_{ih})\lambda_i \lambda_h = K_{hi} - K_{ih}, \quad i, h = 1, \dots, 2k.$$

Hence  $K_{hi} - K_{ih} = 0$  for all  $i$  and  $h$  for which  $\lambda_i \lambda_h \neq 1$ . If the  $\lambda$ 's are distinct, it follows from (7.2) and (7.4) that all the  $K_{ih}$  are zero except the  $K_{2i, 2i-1}$ , ( $i = 1, \dots, k$ ). Thus we find that the differential form (7.3) is equivalent to

$$(7.5) \quad \sum_{i=1}^k M_i x'_{2i} dx'_{2i-1}, \quad M_i = K_{2i, 2i-1}.$$

None of these  $M_i$  can vanish since we are dealing with a nondegenerate Pfaffian transformation. We now consider the following two transformations for changing the variables in  $T$ :

$$(7.6) \quad x''_{2i-1} = x'_{2i-1}, \quad x''_{2i} = M_i x'_{2i},$$

$$(7.7) \quad x'''_{2i-1} = M_i^{1/2} (-1)^{-1/4} x'_{2i-1}, \quad x'''_{2i} = M_i^{1/2} (-1)^{-1/4} x'_{2i},$$

$i = 1, \dots, k.$

Neither of these transformations changes the form of the linear terms in (7.1), while the linear differential form (7.5) appears as  $\sum_{i=1}^k x''_{2i} dx''_{2i-1}$  and  $(-1)^{1/2} \sum x'''_{2i} dx'''_{2i-1}$ .

In case none of the  $\lambda$ 's is in absolute value equal to 1, the notation can

---

\* The reader will find it instructive to consider the various cases arising for  $n=4$  with regard to the four characteristic multipliers  $[\lambda_1, \lambda_2, \lambda_3, \lambda_4]$ . If  $\rho, \sigma, \theta, \phi$  are real, ( $\rho, \sigma \neq 1; \theta, \phi \neq 0, \text{ mod } 2\pi$ ), the four important cases are as follows: I.  $[\rho, \rho^{-1}, \sigma, \sigma^{-1}]$ , II.  $[\rho, \rho^{-1}, e^{\theta(-1)^{1/2}}, e^{-\theta(-1)^{1/2}}]$ , III.  $[e^{\phi(-1)^{1/2}}, e^{-\phi(-1)^{1/2}}, e^{\theta(-1)^{1/2}}, e^{-\theta(-1)^{1/2}}]$ , IV.  $[\rho e^{\theta(-1)^{1/2}}, \rho^{-1} e^{-\theta(-1)^{1/2}}, \rho e^{-\theta(-1)^{1/2}}, \rho^{-1} e^{\theta(-1)^{1/2}}]$ . Case II is the one case in which canonical variables cannot be taken having the property  $R$ . It may be mentioned in this connection that the theorem in Birkhoff [3, p. 78, line 9] is incorrect as may be shown by simple examples.

be chosen in such a way that both members of any pair of imaginary roots will have indices with the same parity; so that, for example,  $\lambda_{2i}$  could be conjugate to  $\lambda_{2j}$  but not to  $\lambda_{2j-1}$ . If this convention is granted it can be shown that  $M_i$  is real, if  $\lambda_i$  is real, but is conjugate to  $M_j$ , if  $\lambda_i$  is conjugate to  $\lambda_j$ , and hence the variables ( $x''$ ) have the property  $R$ .

If all the  $\lambda$ 's are equal in absolute value to 1, then each  $\lambda_{2i}$  is conjugate to its reciprocal  $\lambda_{2i-1}$  and to none other, since we are assuming the  $\lambda$ 's to be distinct. It can then be shown that  $M_i$  is pure imaginary. Let  $M_i = N_i(-1)^{1/2}$ . Without loss of generality we assume  $N_i > 0$ . Otherwise we interchange  $\lambda_{2i-1}$  and  $\lambda_{2i}$ , thus inducing the interchange  $x''_{2i-1}$  and  $x''_{2i}$ . In fact

$$x_{2i}dx_{2i-1} = -x_{2i-1}dx_{2i} + \text{the exact differential } d(x_{2i}x_{2i-1}).$$

With these changes made, if necessary, it is not difficult to show that the variables ( $x'''$ ) have the property  $R$ .

The case of equal latent roots (but simple elementary divisors) may be taken care of by a limiting process.

**8. Normalization of  $Q$  in the formal differential equations.** Let  $T = S^N$ , where  $S$  and  $N$  have the meanings explained at the beginning of §5. Suppose also that  $T$  is a contact transformation and that a preliminary normalization of the linear terms has been carried out as described in §7, so that  $T$  is Pfaffian with respect to  $\sum_{i=1}^k x_{2i}dx_{2i-1}$  and is defined by convergent series of the type

$$(8.1) \quad \bar{x}_{2i-1} = e^{\rho_i x_{2i-1}} + \dots, \quad \bar{x}_{2i} = e^{-\rho_i x_{2i}} + \dots, \\ i = 1, 2, \dots, k = n/2,$$

where only the linear terms have been written. It is assumed also that *there is a positive integer  $P (>3)$  such that there exist no relations of the form  $\sum_{i=1}^k m_i \rho_i = 0$ , where the integers  $m_1, \dots, m_k$  are not all zero and where each  $m_i$  is numerically less than  $P$* . Referring back to the formal Hamiltonian differential equations (6.2), let us write the formal series for  $-Q$  in the form  $-Q = \sum_{\alpha=2}^{\infty} H_{\alpha}$ , where  $H_{\alpha}$  is a homogeneous polynomial of the  $\alpha$ th degree in  $x_1, \dots, x_n$ . The formal series for  $T'$  have linear terms of the same form as those displayed in (8.1) except that  $\rho_i$  is replaced by  $\rho_i t$ . It follows at once from (6.2) that

$$H_2 = \sum_{i=1}^k \rho_i x_{2i-1} x_{2i}.$$

It is now our purpose to make changes in the variables with the help of contact transformations to see if  $H_3, H_4, \dots$  can also be reduced to an especially simple form. Following Birkhoff's treatment of the analogous dynamical

problem,\* we shall apply contact transformations of the form (6.3) with

$$W = \sum_{i=1}^m x_{2i}y_{2i-1} + W_\alpha, \quad \alpha = 3, 4, \dots,$$

where  $W_\alpha$  is a homogeneous polynomial in the  $x_{2i}$  and  $y_{2i-1}$  of degree  $\alpha$ . If we solve explicitly for the  $x$ 's in terms of the  $y$ 's we clearly obtain up to terms of the  $\alpha$ th degree.

$$x_{2i-1} = y_{2i-1} + \frac{\partial W'_\alpha}{\partial y_{2i}} + \dots, \quad x_{2i} = y_{2i} - \frac{\partial W'_\alpha}{\partial y_{2i-1}} + \dots, \quad i = 1, \dots, k.$$

Here  $W'_\alpha$  denotes the function obtained by replacing  $x_{2i}$  by  $y_{2i}$  in  $W_\alpha$ . To terms of the  $\alpha$ th degree inclusive we find therefore that

$$-Q = \sum_{i=1}^k \rho_i y_{2i-1} y_{2i} + \sum_{i=1}^k \rho_i \left[ y_{2i} \frac{\partial W'_\alpha}{\partial y_{2i}} - y_{2i-1} \frac{\partial W'_\alpha}{\partial y_{2i-1}} \right] + \sum_{s=3}^\alpha \bar{H}_s + \dots,$$

where  $\bar{H}_s$  is the polynomial obtained by replacing the  $x$ 's by the  $y$ 's in  $H_s$ .

In order to simplify the terminology let us now change the  $y$ 's into  $x$ 's. Thus, a contact transformation of this type leaves  $H_2, H_3, \dots, H_{\alpha-1}$  unmodified while  $H_\alpha$  takes the form

$$\sum_{i=1}^k \rho_i \left[ x_{2i} \frac{\partial W''_\alpha}{\partial x_{2i}} - x_{2i-1} \frac{\partial W''_\alpha}{\partial x_{2i-1}} \right] + \text{the original } H_\alpha.$$

Here  $W''_\alpha$  represents the function obtained by replacing  $y_{2i-1}$  by  $x_{2i-1}$  in  $W_\alpha$ . Now any term in  $W''_\alpha$  may be written  $c x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$ , the  $\beta$ 's being positive integers whose sum is  $\alpha$ . The corresponding term in the modified  $H_\alpha$  has a coefficient

$$c \left[ \sum_{i=1}^k \rho_i (\beta_{2i} - \beta_{2i-1}) \right] + h,$$

where  $h$  is the coefficient corresponding to  $c$  in the original  $H_\alpha$ . Now, if  $\alpha < P$ , then  $\sum_{i=1}^k (\beta_{2i} - \beta_{2i-1}) \rho_i$  cannot vanish according to the hypothesis italicized above, unless  $\beta_{2i-1} = \beta_{2i}$  for each  $i$ , ( $i = 1, 2, \dots, k$ ). Hence it is possible to choose the coefficients  $c$  in  $W''_\alpha$  (or  $W_\alpha$ ) in such a way that all the terms in the modified  $H_\alpha$  disappear except those which contain  $x_{2i-1}$  and  $x_{2i}$  both to the same power for each  $i$ , at least if  $\alpha < P$ .

Thus carrying out this process successively for  $\alpha = 3, 4, \dots$ , we see that after performing a finite number of contact transformations we may write

$$(8.2) \quad -Q = F(u) + Q_P,$$

\* Birkhoff [3, pp. 82-85].

where  $F(u)$  is a polynomial in the  $k$  products  $x_{2i-1}x_{2i} = u_i$ , ( $i = 1, \dots, k$ ), of degree not greater than  $P/2$ , the linear terms being  $\sum_{i=1}^k \rho_i u_i$ , and where  $Q_P$  is a formal power series in  $x_1, \dots, x_n$  beginning with terms of degree not lower than  $P$ .

The sequence of contact transformations necessary to effect this reduction can, of course, be combined into a single analytic contact transformation. If there are no commensurability relations connecting  $\rho_1, \dots, \rho_k$ ,  $P$  may, of course, be taken arbitrarily large. The process may then be continued indefinitely and we see that there exists a *formal* contact transformation which enables us to write  $Q$  as a formal power series in the  $u$ 's.

9. **Normalized forms for  $T$ .** The above normalization of  $Q$  yields immediately certain normalized forms for nondegenerate Pfaffian transformations. Namely, retaining the hypotheses and notation of the preceding section,  $T$  may be written in the form

$$(9.1) \quad \begin{aligned} \bar{x}_{2i-1} &= x_{2i-1} \exp [\partial F / \partial u_i] + \Xi_{2i-1}, \\ \bar{x}_{2i} &= x_{2i} \exp [-\partial F / \partial u_i] + \Xi_{2i}, \quad i = 1, \dots, k = n/2, \end{aligned}$$

where the  $\Xi$ 's are power series in the  $x$ 's beginning with terms of degree not lower than  $P-1$ . This is proved from the obvious fact that the formal differential equations determine the transformation  $T$  uniquely. Furthermore the power series  $\Xi_i$  must converge since the transformation  $T$  was analytic to begin with and only analytic changes of variable have been used. The remaining formal details of the proof are left to the reader.

The equations (9.1) are especially useful in case the characteristic exponents  $\rho_1, \dots, \rho_k$  are pure imaginary. In this case, if we start out with variables having the property  $R$  in the sense of §7, the variables subsequently introduced will also have this property. The proof of this fact is left to the reader. Transformations of this type have been considered by Birkhoff and Lewis.\* In particular it is known that there exist infinitely many periodic point groups in the neighborhood of the origin, at least, if the Hessian determinant  $|\partial^2 F / \partial u_i \partial u_j|$  evaluated at the origin is distinct from zero. It may be noted here that this result does not depend on complete incommensurability of the characteristic exponents, as assumed in the above mentioned work. It is merely necessary that  $P$  be not less than a certain fixed number which may be taken at least as small as  $16k+10$ .†

\* Birkhoff [4]; Lewis [4, 5]. The transformation referred to appears in different notation at the top of page 119 of the first of these papers.

† In Birkhoff [4], it is merely assumed that  $\mu \geq 8n+4$ . The " $n$ " of that paper is our  $k$  and " $\mu$ " =  $P/2-1$ .

## REFERENCES

## G. D. BIRKHOFF

1. *Surface transformations and their dynamical applications*, Acta Mathematica, vol. 43 (1922), pp. 1-119.
2. *Nouvelles recherches sur les systèmes dynamiques*, Memoriae Pontificiae Academiae Scientiae Novi Lyncaei, (3), vol. 1, pp. 85-216.
3. *Dynamical Systems*, American Mathematical Society Colloquium Publications, vol. 9, New York, 1927.
4. With D. C. Lewis, *On the periodic motions near a given periodic motion of a dynamical system*, Annali di Matematica, (4), vol. 12 (1933), pp. 117-133.

## E. GOURSAT

1. *Leçons sur le Problème de Pfaff*, Paris, 1922.

## D. C. LEWIS

1. *On formal power series transformations*. To be published in the Duke Mathematical Journal.
2. *Invariant manifolds near an invariant point of unstable type*, American Journal of Mathematics, vol. 60 (1938), pp. 577-587.
3. *The formal theory of conservative transformations in  $2n$ -dimensional space*, Bulletin of the American Mathematical Society, abstract 41-5-197.
4. *On certain periodic motions of dynamical systems with more than two degrees of freedom*, American Journal of Mathematics, vol. 56 (1934), pp. 25-41.
5. *Sulle oscillazioni periodiche d'un sistema dinamico*, Rendiconti della Reale Accademia Nazionale dei Lincei, (6\*), vol. 19 (1934), pp. 234-237.

## C. C. MACDUFFEE

1. *The Theory of Matrices*, Berlin, 1933.

## E. VON WEBER

1. *Vorlesungen über das Pfaff'sche Problem und die Theorie der partiellen Differentialgleichungen erster Ordnung*, Leipzig, 1900.

## E. T. WHITTAKER

1. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 2d edition, Cambridge, 1917.

## A. WINTNER

1. With E. R. van Kampen, *On the reduction of dynamical systems by means of parametrized invariant relations*, these Transactions, vol. 44 (1938), pp. 168-195.
2. *Three notes on characteristic exponents and equations of variation in celestial mechanics*, American Journal of Mathematics, vol. 53 (1931), pp. 605-625.

CORNELL UNIVERSITY,  
ITHACA, NEW YORK