ALGEBRAIC PROPERTIES OF SELF-ADJOINT SYSTEMS*

BY

DUNHAM JACKSON

The general definition of adjoint systems of boundary conditions associated with ordinary linear differential equations was given by Birkhoff.† In a paper of Bôcher,‡ in which the idea is further developed, there is obtained a condition that a system of the second order be self-adjoint. It is proposed here to extend the discussion of this problem to the case of systems of any order.§ A condition for self-adjointness of the boundary conditions is simply expressed in matrix form, without any requirement that a corresponding property be possessed by the differential equations; the condition gains in symmetry if it is assumed that the differential expression which forms the left-hand member of the given differential equation is itself identical with its adjoint or with the negative of its adjoint.

As a preliminary, it will be well to recall a well-known rule for the combination of matrices, which will be found particularly convenient. Let α_1, α_2, α_3, and α_4 be n-rowed square matrices. The square matrix of order $2n$ which is made up of the elements of these four, with the elements of α_1, arranged in order, in its upper left-hand corner, and the elements of the other matrices correspondingly disposed, may be indicated by the notation

$$\alpha = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix}.$$

* Presented to the Society, December 27, 1915.

† G. D. Birkhoff, these Transactions, vol. 9 (1908), pp. 373–395; p. 375.

‡ M. Bôcher, these Transactions, vol. 14 (1913), pp. 403–420; p. 409. Cf. E. Bounitzky, Journal de mathématiques pures et appliquées, 6th series, vol. 5 (1909), pp. 65–125; pp. 112–116; W. Stekloff, *Sur certaines questions d'analyse qui se rattachent à plusieurs problèmes de la physique mathématique*, Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg, classe physico-mathématique, 8th series, vol. 31 (1913), no. 7; pp. 5–6, 69.

§ Bounitzky, in the paper just cited, pp. 107–116, gives a discussion of the problem, laying stress primarily on the case of a self-adjoint differential equation. His method of treatment is different from that employed here, and in fact bears as little resemblance to it as could well be expected, considering that the author's task in either case, as far as this particular problem is concerned, is merely to give a concise account of algebraic processes which in themselves are perfectly straightforward.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Let β be another matrix of the same sort, consisting of the elements of the n-rowed matrices $\beta_1, \cdots, \beta_4$. The notation suggests the symbolic product

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} \begin{pmatrix} \beta_1 & \beta_2 \\ \beta_3 & \beta_4 \end{pmatrix} = \begin{pmatrix} \alpha_1 \beta_1 + \alpha_2 \beta_3 & \alpha_1 \beta_2 + \alpha_2 \beta_4 \\ \alpha_3 \beta_1 + \alpha_4 \beta_3 & \alpha_3 \beta_2 + \alpha_4 \beta_4 \end{pmatrix},$$

formed as if the elements indicated were ordinary numbers instead of matrices. By actually writing out the n-rowed matrices $\alpha_1 \beta_1 + \alpha_2 \beta_3$, etc., it is possible to interpret the product as a square matrix of order $2n$; and it is found that this is the same as the product obtained by multiplying together α and β as $2n$-rowed matrices in the usual way.

To begin the discussion of the problem in hand, let $L(u)$ stand for the differential expression

$$L(u) = p_0(x) \frac{d^n u}{dx^n} + p_1(x) \frac{d^{n-1} u}{dx^{n-1}} + \cdots + p_n(x) u,$$

where n is any positive integer. The coefficients p are real or complex functions of the real variable x, defined in an interval $a \leqq x \leqq b$, and it may be assumed, with more regard for simplicity of statement than for ultimate generality, that they are continuous together with their first n derivatives. With the differential equation

$$L(u) = 0$$

there is supposed given a set of n linearly independent boundary conditions $U_i(u) = 0$, $i = 1, 2, \cdots, n$, in which the general left-hand member is a linear combination, with real or complex constant coefficients, of the $2n$ quantities $u(a), u'(a), \cdots, u^{(n-1)}(a), u(b), u'(b), \cdots, u^{(n-1)}(b)$. Let n of these quantities be chosen in such a way that the determinant of the corresponding coefficients is different from zero. The possibility of such a choice is a consequence of the assumption that the n conditions are linearly independent The quantities selected will be denoted by $u_1, u_2, \cdots, u_n$, the remaining n quantities by $u_{n+1}, \cdots, u_{2n}$; the order of the subscripts within each group is immaterial. It is most natural to assign the first n subscripts to $u(a), \cdots, u^{(n-1)}(a)$, or to $u(b), \cdots, u^{(n-1)}(b)$, if the corresponding sets of terms are linearly independent, but this will not always be the case.

The differential expression adjoint to $L(u)$ is

$$M(v) = (-1)^n \frac{d^n}{dx^n}(p_0 v) + (-1)^{n-1} \frac{d^{n-1}}{dx^{n-1}}(p_1 v) + \cdots + p_n v.$$

It is connected with $L(u)$ by Lagrange's identity,

$$vL(u) - uM(v) = \frac{d}{dx} P(u, v),$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

where $P(u, v)$ is a bilinear form in $u(x)$, $u'(x)$, $\cdots$, $u^{(n-1)}(x)$; $v(x)$, $v'(x)$, $\cdots$, $v^{(n-1)}(x)$, with coefficients which are functions of x. Integrated with regard to x from a to b, this identity gives a relation involving the values which the functions just named take on at the ends of the interval. If the values of v and its derivatives at a and b are denoted by v_1, $\cdots$, v_{2n}, in such a way that u_j and v_j refer to derivatives of the same order taken at the same point, the relation becomes

$$\int_a^b [vL(u) - uM(v)]\,dx = \Pi(u, v),$$

in which $\Pi(u, v)$ is a non-singular bilinear form in the $4n$ quantities u_j, v_j, with constant coefficients.*

The definition of the adjoint boundary conditions proceeds now as follows. In addition to the linear forms $U_1(u)$, $\cdots$, $U_n(u)$, there are defined n other forms $U_{n+1}(u)$, $\cdots$, $U_{2n}(u)$, which are arbitrary, except for the restriction that the whole set of $2n$ forms shall be linearly independent. Then $2n$ linear forms $V_1(v)$, $\cdots$, $V_{2n}(v)$, in the variables v_j, are determined so that

$$\Pi(u, v) = \sum_{i=1}^{2n} U_i(u)\,V_i(v).$$

The last n of these forms, set equal to zero, express the adjoint conditions. It is well known that a different choice of the arbitrary forms U_{n+1}, $\cdots$, U_{2n}, would lead to a set of conditions equivalent collectively to those already obtained, that is, a set of independent linear combinations of them.† It is to be kept in mind throughout that it is the set of equations that is important, and not the individual equation, still less the individual linear form.

The definition just given can be expressed in slightly different words, as follows: The variables u are subjected to a linear transformation $u'_i = U_i(u)$, and then a linear transformation of the other set of variables, $v'_i = V_i(v)$, is determined so that the two transformations together carry over the bilinear form $\Pi\ (u, v)$ into the normal form $\sum u'_i\, v'_i$. Let Π denote the matrix of the coefficients in the form $\Pi(u, v)$; the double use of the letter Π, to represent a matrix and to represent an operator, will not cause any confusion. Let α denote the matrix of the coefficients in the *inverse* transformation, by which the u's without accents are expressed in terms of the u's with accents, and let β be the corresponding matrix for the v's; the conjugate of α will be de-

* It is to be noticed that the coefficient p_n is without effect on the form of Π, as the term $p_n\,uv$ cancels from the difference $vL(u) - uM(v)$. Consequently, if the differential equation with a parameter, $L(u) + \lambda u = 0$, is under discussion, the form Π is independent of the parameter, and the same is true of the adjoint boundary conditions, which depend on Π for their definition.

† Birkhoff, loc. cit., p. 375; Bôcher, loc. cit., p. 405.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

noted by α', and, in general, conjugate matrices will be indicated by accents. Finally, let I denote the unit matrix of order $2n$, having unity for each element in the principal diagonal, and zero everywhere else. Then the relation connecting these matrices is simply the equation*

$$\alpha' \Pi \beta = I. \tag{1}$$

It is to be noticed that not α, but α^{-1}, is to be regarded as given directly, and β^{-1} is the matrix that is ultimately desired. An expression for the latter is found by multiplying both members of (1) on the right by β^{-1}:

$$\beta^{-1} = \alpha' \Pi.$$

Hence the problem to be solved may be formulated thus:

Under what circumstances will the last n rows of $\alpha' \Pi$ be independent linear combinations of the first n rows of α^{-1}?

Let α_1 stand for the square matrix of the coefficients of the first n variables u_j in the n given forms U_i, and α_2 for the matrix of the remaining coefficients in these forms. Since the notation has been supposed chosen so that α_1 is non-singular, the $2n$ forms $U_1, \cdots, U_{2n}$ will be linearly independent if the last n of them are set equal to $u_{n+1}, \cdots, u_{2n}$ respectively. Then the matrix α^{-1} has the form

$$\alpha^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & I \end{pmatrix};$$

here I denotes the unit matrix of order n (the use of the same letter to represent unit matrices of different orders will not cause any confusion) and 0 stands for a matrix composed entirely of zeros. It may be verified at once that

$$\begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & I \end{pmatrix} \begin{pmatrix} \alpha_1^{-1} & -\alpha_1^{-1}\alpha_2 \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

from which it follows that the second factor on the left is the matrix α. The product $\alpha_1^{-1}\alpha_2$ will be denoted for brevity by the single letter δ. It will be convenient also to indicate the conjugate of the reciprocal of a non-singular matrix by a double accent, so that $(\alpha_1^{-1})' = (\alpha_1')^{-1} = \alpha_1''$, and $\delta' = \alpha_2' \alpha_1''$. Then we have for the conjugate of α the representation

$$\alpha' = \begin{pmatrix} \alpha_1'' & 0 \\ -\delta' & I \end{pmatrix}.$$

The matrix of the bilinear form will be represented for the present simply by the notation

$$\Pi = \begin{pmatrix} \pi_1 & \pi_3 \\ \pi_4 & \pi_2 \end{pmatrix},$$

* Cf., e. g., Bôcher, *Introduction to Higher Algebra*, p. 115.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

which is applicable to any square matrix of order $2n$; with regard to the special form of Π, more will be said later. The expression for β^{-1} is

$$\alpha' \Pi = \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_3 \\ -\delta' \pi_1 + \pi_4 & -\delta' \pi_3 + \pi_2 \end{pmatrix}.$$

Now, since α_1 is non-singular, it is possible to express any sets of n quantities each, in particular, the rows of the matrix standing in the lower left-hand corner of $\alpha' \Pi$, as linear combinations of the rows of α_1; and the coefficients by which this is accomplished are uniquely determined. The question is, then, whether the corresponding linear combinations of the rows of α_2 are identical respectively with the rows of the matrix in the lower right-hand corner of $\alpha' \Pi$. It will be seen presently that if this is the case, the combinations in question will necessarily be independent.

In general terms, if η is a matrix of order n, and there are formed n linear combinations of its rows, the coefficients in these combinations forming the successive rows of a matrix γ, the resulting sets of n quantities each are respectively the rows of the product matrix $\gamma\eta$. In the present instance, γ is to be determined so that

$$\gamma\alpha_1 = -\delta' \pi_1 + \pi_4, \tag{2}$$

and the condition for a self-adjoint system is that the matrix γ so defined satisfy the further relation

$$\gamma\alpha_2 = -\delta' \pi_3 + \pi_2. \tag{3}$$

If it does, then

$$\alpha' \Pi = \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_3 \\ \gamma\alpha_1 & \gamma\alpha_2 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} \alpha_1'' \pi_1 & \alpha_1'' \pi_2 \\ \alpha_1 & \alpha_2 \end{pmatrix}.$$

If γ were singular, the matrix $\alpha' \Pi$ would be singular, which is certainly not the case; it follows that the determinant of γ is necessarily different from zero. Let (2) be solved for γ, and the result substituted in (3). It is found that

$$\gamma = -\delta' \pi_1\alpha_1^{-1} + \pi_4 \alpha_1^{-1}, \qquad -\delta' \pi_1 \delta + \pi_4 \delta = -\delta' \pi_3 + \pi_2. \tag{4}$$

In the last equation, δ has been substituted in the left-hand member for its equal, $\alpha_1^{-1} \alpha_2$. The relation (2), defining γ, may be regarded as valid in any case. Then either of the relations (3), (4), is a consequence of the other. The conclusion may be stated as follows:

THEOREM. *A necessary and sufficient condition that the given set of boundary conditions be self-adjoint, is that**

$$\delta' \pi_1 \delta - \pi_4 \delta = \delta' \pi_3 - \pi_2. \tag{5}$$

* Professor W. A. Hurwitz points out that this condition can be written concisely in the form $\Delta' \Pi \Delta = 0$, where

$$\Delta = \begin{pmatrix} 0 & -\delta \\ 0 & I \end{pmatrix}.$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

In the cases that are of most interest, the form of the condition can be simplified. Suppose first that in the n given boundary conditions, the sets of terms corresponding to one end of the interval are linearly independent, so that the values of u and its derivatives at this point may be represented by $u_1, \cdots, u_n$. Then, as no single term of $\Pi(u, v)$ is concerned with both ends of the interval, the subscripts in each term that actually appears will be both less than $n + 1$ or both greater than n, and the matrices previously denoted by π_3 and π_4 will be composed entirely of zeros.* If δ is replaced by the alternative expression $\alpha_1^{-1} \alpha_2$, the condition for self-adjointness reads as follows:

$$\alpha_2' \alpha_1'' \pi_1 \alpha_1^{-1} \alpha_2 = - \pi_2 .$$

It appears that in this case α_2 as well as α_1 must be non-singular, since π_2 is non-singular. On multiplication before and after by α_2'' and α_2^{-1} respectively, the condition takes the form

$$\alpha_1'' \pi_1 \alpha_1^{-1} = - \alpha_2'' \pi_2 \alpha_2^{-1} ,$$

the meaning of which can be expressed in still different language. There is no loss of generality in assuming, for definiteness, that α_1 is the matrix of the coefficients belonging to the point b. More precisely, it may be assumed that

$$u_1 = u(b), \qquad u_2 = u'(b), \qquad \cdots, \qquad u_n = u^{(n-1)}(b),$$

$$u_{n+1} = u(a), \qquad u_{n+2} = u'(a), \qquad \cdots, \qquad u_{2n} = u^{(n-1)}(a),$$

while $v_1, \cdots, v_{2n}$ have a corresponding interpretation. Let the value of $P(u, v)$ for $x = b$ be denoted by $\Pi_1(u, v)$, and its value for $x = a$, by $\Pi_2(u, v)$, so that

$$\Pi(u, v) = \Pi_1(u, v) - \Pi_2(u, v) .$$

The expressions Π_1 and Π_2 are bilinear forms in $u_1, \cdots, u_n; v_1, \cdots, v_n$, and $u_{n+1}, \cdots, u_{2n}; v_{n+1}, \cdots, v_{2n}$, respectively, and their matrices are π_1 and $- \pi_2$. Let each of the n given linear forms $U_i(u)$ be regarded as the sum of a form $U_{i1}(u)$, involving only the first n of the variables u_j, and a form $U_{i2}(u)$, involving only the last n variables. If $\Pi_1(u, v)$ is written as a bilinear form in $U_{i1}(u)$ and $U_{i1}(v)$, $i = 1, 2, \cdots, n$, the coefficients in the representation have the matrix $\alpha_1'' \pi_1 \alpha_1^{-1}$. A corresponding remark applies to $\Pi_2(u, v)$. It follows that

If the forms $U_{i1}(u)$ are linearly independent, the condition for a self-adjoint system is that the coefficients by means of which $\Pi_1(u, v)$ is expressed in terms of the linear forms $U_{i1}(u)$, $U_{i1}(v)$, be the same as those by which $\Pi_2(u, v)$ is expressed in terms of the forms $U_{i2}(u)$, $U_{i2}(v)$.

---

* The fact that π_1 and π_2 also contain a considerable number of zeros need not be taken into account.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

Now let the restriction on the choice of the first n variables be removed, but let it be supposed that when the given differential expression $L(u)$ and its adjoint are written with the same function u as argument, the latter expression is identical with $L(u)$ or with the negative of $L(u)$, in other words, that $L(u)$ is self-adjoint or anti-self-adjoint. The statement of the theorem is again capable of some simplification, for the reason now that the bilinear form $P(u, v)$ is skew-symmetric in the one case and symmetric in the other,* and the same thing consequently is true of the matrix Π. More generally, it suffices for the symmetry of P and Π that $L(u)$ differ only in the term $p_n u$ from an expression which is anti-self-adjoint,† since, as has already been remarked, this term has no effect on the form of P.

Let the condition (5) be written in the equivalent form

$$\delta' \pi_1 \delta - 2\pi_4 \delta + \pi_2 = - \delta' \pi_1 \delta + 2\delta' \pi_3 - \pi_2. \tag{6}$$

If Π is skew-symmetric, then $\pi_1' = -\pi_1$, $\pi_2' = -\pi_2$, and $\pi_4' = -\pi_3$, and it appears that the matrix which forms the right-hand member of (6) is the conjugate of the matrix on the left. If Π is symmetric, each member of (6) is the negative of the conjugate of the other. That is:

If the differential expression $L(u)$ is self-adjoint, the condition that the boundary conditions be self-adjoint is that the matrix

$$\delta' \pi_1 \delta - 2\pi_4 \delta + \pi_2$$

be symmetric; if $L(u)$ is an anti-self-adjoint expression, or differs from such an expression only in the term of order zero, the condition is that the matrix just written down be skew-symmetric.

This form of statement remains valid, of course, under the earlier hypothesis that $\pi_4 = 0$; but in that case the condition simply resumes the form

$$\delta' \pi_1 \delta + \pi_2 = 0,$$

since this matrix is required to be symmetric and skew-symmetric at the same time.‡

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

* Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. 1, pp. 70–71, 73–74.

† This observation is not trivial, as the corresponding statement with regard to self-adjoint expressions would be; for example, the expression $xu' + \frac{1}{2}u$ is anti-self-adjoint, while the expression xu' is not.

‡ It is necessarily skew-symmetric if $L(u)$ is self-adjoint, without regard to the boundary conditions, and becomes symmetric if the latter are self-adjoint, and similarly in the anti-self-adjoint case.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use