A generalization of the Sturm[1]) Separation and Comparison Theorems in n-space.

Von

Marston Morse in Cambridge (Mass., U.S.A.).

§ 1.
Introduction.

The author considers a system of n ordinary, second-order, linear, homogeneous differential equations. The system is supposed self-adjoint and regular. Starting with the theorem that any such system may be considered as the Euler equations of an integral of a quadratic form, he shows that the Sturm Separation and Comparison Theorems naturally generalize into theorems concerning the interrelation of conjugate and focal points. Care is taken at every stage to state the final results in a purely analytical form belonging to the theory of differential equations proper.

The results obtained were made possible by certain other discoveries already made in connection with the author's *theory of the calculus of variations in the large*[2]).

§ 2.
Self-adjoint systems.

Let there be given a system of n differential equations of the form[3])

$$(2.1)\qquad a_{ij} z_j'' + b_{ij} z_j' + c_{ij} z_j = 0 \qquad (i, j = 1, \ldots, n), \ |a_{ij}| \neq 0$$

---

[1]) See Bôcher, Leçons sur les Méthodes de Sturm, Paris (1917), Gauthier-Villars et Cie.

[2]) Marston Morse, The foundations of the calculus of variations in the large in m-space (first paper). Transactions of the American Mathematical Society **31**, No. 3 (1929). See also, The critical points of functions and the calculus of variations in the large, Bulletin of the American Mathematical Society **35** (1929), pp. 38–54. The first paper will hereafter be referred to as Morse.

[3]) We adopt the convention that a repetition of a subscript in a term shall indicate a sum with respect to that subscript.

where x is the independent variable, and a_{ij}, b_{ij}, and c_{ij}, are of class C'' on a segment γ of the x axis. The system adjoint to (2.1) is

$$(2.2) \qquad (a_{ij} y_i)'' - (b_{ij} y_i)' + (c_{ij} y_i) = 0.$$

A set of necessary and sufficient conditions that (2.1) be self-adjoint is readily seen to be[4])

$$(2.3)' \qquad a_{ij} \equiv a_{ji},$$

$$(2.3)'' \qquad 2\,a'_{ij} \equiv b_{ij} + b_{ji},$$

$$(2.3)''' \qquad a''_{i} - b'_{ij} \equiv c_{ji} - c_{ij}.$$

On the other hand let us turn to the integral

$$\mathrm{I} = \int_{x_0}^{x_1} (A_{ij} z'_i z'_j + 2\,B_{ij} z_i z'_j + C_{ij} z_i z_j)\,dx$$

where A_{ij}, and B_{ij}, and C_{ij} are functions of class C' on γ. No loss of generality results if we assume

$$(2.4) \qquad A_{ij} \equiv A_{ji}, \qquad C_{ij} \equiv C_{ji}.$$

The Euler equations corresponding to the integral I are

$$(2.5) \qquad \frac{d}{dx}(A_{ij} z'_j + B_{ji} z_j) - (B_{ij} z'_j + C_{ij} z_j) = 0.$$

We see that a set of necessary and sufficient conditions for the system (2.5), when expanded, to be identical with the system (2.1), is the following

$$(2.6)' \qquad a_{ij} \equiv A_{ij},$$

$$(2.6)'' \qquad b_{ij} \equiv A'_{ij} + B_{ji} - B_{ij},$$

$$(2.6)''' \qquad c_{ij} \equiv B'_{ji} - C_{ij}.$$

We shall prove the following theorem.

Theorem 1. *A necessary and sufficient condition that the system* (2.1) *admit the form* (2.5) *is that the system* (2.1) *be self-adjoint.*

That it is necessary that (2.1) be self-adjoint is a well known fact[5]) whose proof we need not give.

To prove that it is sufficient that (2.1) be self-adjoint for (2.1) to admit the form (2.5), we assume that the equations (2.3) hold, and

[4]) See Davis R. D., The inverse problem of the calculus of variations in higher space, Transactions of the American Mathematical Society **30** (1928), p. 711–713.

[5]) See Hadamard, Leçons sur le calcul des variations, Paris 1910, **1**, p. 319.

then show that it is possible to choose A_{ij}, B_{ij}, and C_{ij}, so as to satisfy (2.4) and (2.6).

We first choose A_{ij} as a_{ij}. We then choose the functions C_{ij} as arbitrary functions of class C' except for the conditions $C_{ij} = C_{ji}$. We next choose B_{ji} and B_{ij} so as to satisfy $(2.6)''$ as an initial condition at some point $x = x_0$. Equations $(2.6)''$ give apparently two conditions on $B_{ji} - B_{ij}$, obtained by interchanging i and j, namely,

$$B_{ji} - B_{ij} \equiv b_{ij} - A'_{ij},$$
$$B_{ij} - B_{ji} \equiv b_{ji} - A'_{ji}.$$

But these two conditions are really the same as is seen from the fact that A_{ij} has been chosen as a_{ij}, and from the fact that $(2.3)''$ is assumed to hold. Subject to this condition we now choose B_{ji} and B_{ij} so as to satisfy $(2.6)'''$.

We have made our choices of A_{ij}, B_{ij}, and C_{ij}. It remains to show that $(2.6)''$ holds not only at $x = x_0$ but identically.

According to the choices of B_{ij} and B_{ji} as satisfying $(2.6)'''$, we have

$$B'_{ji} - B'_{ij} \equiv c_{ij} - c_{ji}.$$

By virtue of (2.3) and of our choice of A_{ij} as a_{ij} this becomes

$$A''_{ij} + B'_{ji} - B'_{ij} \equiv b'_{ij}. \tag{2.7}$$

Equation (2.7) is also obtainable by differentiating $(2.6)''$. Since B_{ij} and B_{ji} have been chosen so that $(2.6)''$ holds for at least one x, by virtue of (2.7) it holds identically.

Thus the system (2.6) is satisfied by our choices of A_{ij}, B_{ij}, and C_{ij}, and the theorem is proved.

We note the following.

We can choose $C_{ij} = C_{ji}$ arbitrarily, and $A_{ij} = a_{ij}$. The functions B_{ii} are then uniquely determined to an arbitrary constant. The functions B_{ij} for $i \neq j$, are uniquely determined, except that an arbitrary constant may be added to an admissible choice of B_{ij} if substracted from the corresponding admissible choice of B_{ji}.

The second variation of the integral I takes the form of $2I$ with (z) replaced by (η) as is conventional. Thus if a system (2.1) is self-adjoint it may be considered as the system of Jacobi differential equations of a problem of the calculus of variations. Conversely, as is well known, the Jacobi differential equations of an ordinary problem in non-parametric form will always reduce to a system (2.1) that is self-adjoint.

§ 3.

A canonical form for the differential equations.

We now make a transformation from the variables (z) to variables (y) of the form

$$z_i = u_{ik}(x)\, y_k \qquad (i, k = 1, \ldots, n). \tag{3.1}$$

We seek to determine the functions $u_{ij}(x)$ so that they shall be of class C', their determinant shall not be zero, and so that after (3.1) the terms in $y_h y'_k$ shall disappear from the integral I.

We can also write (3.1) in the form

$$z_j = u_{jh}\, y_h \qquad (j, h = 1, \ldots, n). \tag{3.2}$$

If we make these substitutions in the integral I the terms in the integrand involving both the y's and their derivatives take the form

$$A_{ij} u'_{ik} u_{jh} y_k y'_h + A_{ij} u_{ik} u'_{jh} y'_k y_h + 2 B_{ij} u_{ik} u_{jh} y_k y'_h. \tag{3.3}$$

The second sum in (3.3) will not be changed if we interchange h with k and i with j. The sum (3.3) then becomes

$$2\,[A_{ij} u'_{ik} + B_{ij} u_{ik}]\, u_{jh} y_k y'_h. \tag{3.4}$$

We can make this sum vanish identically by choosing u_{ik} so that

$$A_{ij} u'_{ik} + B_{ij} u_{ik} = 0. \tag{3.5}$$

We choose the kth column of $|u_{ik}|$ as that solution of the n differential equations

$$A_{ij} u'_i + B_{ij} u_i = 0 \qquad (i, j = 1, 2, \ldots, n),$$

which takes on the values

$$u_k(a) = 1, \quad u_i(a) = 0 \qquad (i \neq k).$$

For this choice of the elements u_{ij} the determinant $|u_{ik}|$ is not zero at $x = a$, and is accordingly never zero.

The transformation (3.1) so determined reduces the integral I to the form

$$I = \int_{x_0}^{x_1} \frac{1}{2}\, [r_{ij}(x)\, y'_i y'_j + p_{ij}(x)\, y_i y_j]\, dx \tag{3.6}$$

where

$$r_{ij} = r_{ji}, \quad p_{ij} = p_{ji}, \qquad (|r_{ij}| \neq 0).$$

Our differential equations have thus been reduced to the canonical form,

$$\frac{d}{dx}(r_{ij} y'_j) - p_{ij} y_j = 0. \tag{3.7}$$

We shall assume that the calculus of variations problem arising from I is regular. That is we assume that

$$r_{ij}(x)\eta_i\eta_j > 0 \tag{3.8}$$

for every set $(\eta) \neq 0$, *and for every value of* x *on the given interval.*

§ 4.

Conjugate families of solutions.

From this point on we shall take our system in the form (3.7). If (y) and (z) stand for any two solutions of (3.7) it follows readily from (3.7) that

$$r_{ij}(y_i z_j' - z_i y_j') \equiv \text{const.} \tag{4.1}$$

If the constant in (4.1) is zero (y) and (z) are called *mutually conjugate* according to von Escherich[6]).

A system of n linearly independent mutually conjugate solutions will be called a *conjugate base.* The set of all solutions linearly dependent on the solutions of a conjugate base will be called a *conjugate family.*

By a determinant $D(x)$ of a conjugate base is meant the determinant whose columns are the respective solutions of the base. A determinant $D(x)$ cannot vanish identically. In fact one can easily show that it vanishes at a point $x = a$ at most to a finite order equal to the nullity[7]) of $D(a)$. (Morse § 7.) It is also clear that the determinants of two different conjugate bases of the same conjugate family are non-vanishing constant multiples of each other.

A zero of the rth order of a determinant $D(x)$ will be called a *focal point* of the rth order of the corresponding family.

If $x = a$ is not a focal point of a given conjugate family, one can always select a conjugate base of elements $y_{ij}(x)$ which at a point $x = a$ take on the values

$$y_{ii}(a) = 1, \quad y_{ij}(a) = 0 \qquad (i \neq j). \tag{4.2}$$

Such a base will be termed *unitary* at $x = a$.

§ 5.

The initial elements $g_{ij}(a)$ of a conjugate family.

We shall now investigate with what freedom one can determine a conjugate family by prescribing initial conditions at $x = a$.

---

6) See Bolza, Vorlesungen über Variationsrechnung, p. 626.

7) The nullity of a determinant is its order minus its rank.

We shall restrict ourselves in this section to conjugate families for which $x = a$ is not a focal point. Such a family will have a conjugate base $y_{ij}(x)$ that is unitary at $x = a$. Let us set

$$g_{ij}(a) = r_{ik}(a)\, y'_{kj}(a). \tag{5.1}$$

We see that $g_{ij} = g_{ji}$. For the condition that the ith and jth column of the base $y_{ij}(x)$ be conjugate is that

$$r_{hk}(y_{hi}\, y'_{kj} - y_{hj}\, y'_{ki}) \equiv 0 \tag{5.2}$$

which reduces at $x = a$ with the aid of (4.2) to the condition $g_{ij} = g_{ji}$.

Conversely if the g_{ij}'s are arbitrary elements of a symmetric matrix one sees at once that a matrix of functions $y_{ij}(x)$ whose columns satisfy the differential equations, whose values at $x = a$ are given by (4.2), and which satisfy (5.1), will also satisfy (5.2), and thus be a conjugate base that is unitary at $x = a$.

We have thus proved the following.

Lemma. *The most general conjugate family without focal point at $x = a$ may be determined by giving an arbitrary symmetric matrix of constants $g_{ij}(a)$, and determining a conjugate base that is unitary at $x = a$ and satisfies* (5.1).

When the constants $g_{ij}(a)$ are related to a conjugate family as in the preceding lemma the constants $g_{ij}(a)$ will be called *the initial elements of the family at $x = a$.*

§ 6.

Transverse manifolds.

We need to develop the theory of conjugate families in connection with the theory of transverse manifolds. But the latter theory proceeds with difficulty if the integral be zero along an extremal. Now the integrand H of I is zero along γ. We can avoid any difficulty by taking a new integrand

$$L(x, y, y') = H(x, y, y') + 1. \tag{6.1}$$

The corresponding integral

$$J = \int_{x_0}^{x_1} L(x, y, y')\, dx \tag{6.2}$$

will have the same extremals as the previous one.

Relative to our new integral J we shall now briefly indicate a proof of the following lemma.

Lemma. *Through each point P of γ at which $D(x) \neq 0$ there passes an n-manifold S which in the neighborhood of P is regular*[8]) *and of class C'', and which cuts the extremals of the given conjugate family transversally.*

Let $y_{ij}(x)$ be the elements of a base of the given conjugate family. The extremals of the conjugate family may be represented in the form

$$(6.3) \qquad y_i = c_j y_{ij}(x) \qquad (i, j = 1, 2, \ldots, n).$$

In $L(x, y, y')$ let y_i and y_i' be replaced by the right members of (6.3) and their derivatives respectively. The condition that the hth and kth columns of our base be mutually conjugate may now be written in the form

$$(6.4) \qquad \frac{\partial y_i}{\partial c_h}\frac{\partial L_{y_i'}}{\partial c_k} - \frac{\partial y_i}{\partial c_k}\frac{\partial L_{y_i'}}{\partial c_h} = 0 \qquad (i, h, k = 1, 2, \ldots, n).$$

In this form the condition (6.4) may be identified with the usual condition appearing in the literature[9]) that the Hilbert invariant integral I^0, set up for the family (6.3), be independent of the path, at least in the neighborhood of any point $x = a$ on γ at which $|y_{ij}| \neq 0$. In the neighborhood of $x = a$ on γ the Hilbert integral then becomes a function $I^0(x, y_1, \ldots, y_n)$ of class C''. The equation

$$(6.5) \qquad I^0(x, y_1, \ldots, y_n) = I^0(a, 0, \ldots, 0)$$

is solvable for x as a function $\bar{x}(y)$ of the y_i's, since

$$I_x^0(a, 0) = L(a, 0, 0) = 1 \neq 0.$$

Equation (6.5) thus defines a manifold S which in the neighborhood of $x = a$ on γ satisfies the lemma.

Reference to the form of I^0 shows that all the partial derivatives of I^0 with respect to the y_i's vanish when $(y) = (0)$, from which it follows that S is orthogonal to the x axis at $x = a$.

If now we suppose the base $y_{ij}(x)$ is unitary at $a = x$ we may apply the lemma of Morse § 16 to the jth column of y_{ij}, and thereby obtain the following.

$$\bar{x}_{y_i y_j}(0) + r_{ik}(a) y'_{kj}(a) = 0 \qquad (i, j, k = 1, \ldots, n).$$

If we compare this with (5.1) *we see that*

$$(6.6) \qquad \bar{x}_{y_i y_j}(0) = -g_{ij},$$

a relation that will be useful presently.

---

[8]) An n-manifold S is called regular if on it the coordinates of points neighboring a given point admit a representation in terms of n parameters in which not all of the jacobians of the coordinates with respect to the n parameters are zero.

[9]) Bliss, The transformations of Clebsch in the calculus of variations, Transactions of the American Mathematical Society 17 (1916), p. 595.

§ 7.

The fundamental quadratic form of a conjugate family.

We shall now review certain theorems which show how to characterize focal points in terms of a fundamental quadratic form. Consider a conjugate family *without focal point* at $x = a$ on γ. Let S be the manifold which cuts the family transversally at $x = a$.

Let us cut across the portion of γ for which $a < x < b$ with m successive n-planes t_i, perpendicular to γ, cutting respectively in points at which $x = x_i$. These n-planes divide γ into $m+1$ successive segments. Suppose these n-planes are placed so near together that no one of these $m+1$ segments contains a conjugate point of its initial end point. Let P_i be any point on t_i near γ. Let the points on γ at which $x = a$ and $x = b$ respectively be denoted by A and B. Let Q denote any point near A on the manifold S.

The points

$$Q, P_1, P_2, \ldots, P_m, B \tag{7.1}$$

can be successively joined by extremal segments neighboring γ. Let the resulting broken extremal be denoted by E. Let (v) be a set of $\mu = (m+1)n$ variables of which the first n are the coordinates (y) of Q, the second n are those of P_1, and so on, until finally the last n are the coordinates (y) of P_m. The value of the integral I taken along E will be a function of the variables (v) of class C'' at least, and will be denoted by $J(v)$.

The function $J(v)$ will have a critical point when $(v) = (0)$. We set

$$H(u) = J_{v_h v_k}(0)\, u_h u_k \qquad (h, k = 1, 2, \ldots, \mu). \tag{7.2}$$

The form $H(u)$ will be called the *fundamental form* of the given conjugate family taken from $x = a$ to $x = b$.

Consider now the points

$$P_0, P_1, \ldots, P_m, B, \tag{7.3}$$

where Q in (7.1) is here replaced by a point P_0 in the plane $x = a$. Let (u) be a complex of μ variables composed of the sets of coordinates (y) of the points (7.3) omitting B, and taking these sets in the order of the points (7.3). A curve of class C' which passes from P_0 to B through points (7.3) will be said to *determine* the corresponding set (u).

If we differentiate the function

$$J(e u_1, \ldots, e u_\mu)$$

twice with respect to e, and set $e = 0$, making use of (6.6), we obtain, as in the proof of the lemma, Morse § 17, the following result.

Lemma. *The fundamental form $H(u)$ of the given conjugate family is given as follows:*

$$(7.4) \qquad H(u) = g_{ij} u_i u_j + \int_a^b (r_{ij} \eta_i' \eta_j' + p_{ij} \eta_i \eta_j) dx \qquad (i, j = 1, \ldots, n)$$

where g_{ij} is the ijth initial element of the family at $x = a$, and $[\eta(x)]$ gives the coordinates (y) along the broken extremal from P_0 to B which determines the set (u).

We come to the following theorem.

Theorem 2. *If the point $x = b$ is a focal point of the given family of the rth order, the rank of $H(u)$ is $\mu - r$. If $x = b$ is not a focal point of the family the rank of $H(u)$ is μ, and the negative type number*[10]*) of $H(u)$ equals the number of focal points of the given family between $x = a$ and $x = b$, always counting focal points according to their orders.*

For the reader who has examined the proofs of Theorems 1 to 4 of the earlier paper it will suffice here to enumerate the essential steps as follows.

a) The nullity of the matrix of $H(u)$ will equal the number of linearly independent solutions (u) of the μ equations $H_{u_i} = 0$. The latter equations when expressed in the terms of the right member of (7.4), are seen to be necessary and sufficient conditions on a set (u) for such a set to determine a member of the conjugate family passing through B. If B is a focal point of the family of the rth order there will be exactly r such sets which are linearly independent. The statements of the theorem about the rank of $H(u)$ follow.

b) To determine the type numbers of $H(u)$ we vary the position of the point b from a point nearer $x = a$ than any focal point of the family, to its final position as given. At the start of this variation we see that $H(u)$ will be of negative type zero. As b increases through each focal point the negative type number of $H(u)$ can change by at most the nullity of $H(u)$, that is by at most the order of the focal point. Hence the negative type number is at most the sum q of the orders of the focal points.

c) A lemma on quadratic forms could be stated as follows. A quadratic form which is negative on a q-plane through the origin, excepting the origin, will be of negative type at least q.

---

[10]) By the positive and negative type numbers of a real quadratic form, will be understood the number of positive and negative terms respectively in the form when reduced by a real non-singular linear transformation to squared terms only.

Let $x = c$ be any focal point of the family. We now join B to a point P_0 on the n-plane $x = a$ as follows. We pass along the x axis from B to $x = c$, and then along any curve of the conjugate family to a point P_0 on $x = a$. Let (u) be the set determined by this curve. If c is of order r we can get r such sets which are linearly independent, and from all the focal points we can get q sets (u) which are linearly independent.

Each of the curves just described and their linear combinations will make the right member of (7.4) zero if (η) in that member be taken along such a curve, and (u) be the set which that curve determines. This is proved by suitably integrating the first sum in the integral by parts. If now each of those curves be replaced by the broken extremal which determines the same (u), the integral will be decreased, at least unless some of the points $x = x_i$ are focal points, or unless $(\eta) \equiv (0)$.

Without changing the type numbers of $H(u)$ we can, however, always displace the points $x = x_i$ so that they are not focal points. We conclude that $H(u)$ is negative at each point of a q-plane π_q through the origin, excepting the origin, so that the negative type number of $H(u)$ must be exactly q.

Thus the theorem is proved.

The preceding proof also includes a proof of the following lemma.

Lemma A. *Suppose none of the points x_i are focal points but that $x = b$ is a focal point of the rth order. Let q equal the sum of the orders of the focal points on the interval $a < x \leqq b$. Then there exists a q-plane π_q through the origin, at each point of which $H(u) < 0$ except at the points of a sub r-plane π_r of points (u) determined by the members of the conjugate family through B.*

At no point of π_q excepting the origin are the first n of the u_i's all zero.

The last statement follows from the fact that the only member of the conjugate family which vanishes at $x = a$ is coincident with the x axis.

§ 8.

The special quadratic form for conjugate points.

Suppose now that we have a conjugate family every member of which vanishes at $x = a$. The focal points of such a family determine what are called the conjugate points of $x = a$. We proceed with this case much as in § 7, defining the n-planes t_i, the points x_i, the points P_i, A, and B.

The points

$$(8.1) \qquad A, P_1, \ldots, P_m, B$$

determine a broken extremal E along which I is evaluated.

By the set (u) is here meant the $\nu = mn$ variables of which the first n are the y-coordinates of P_1, the second n those of P_2, etc., and the last those of P_m. The integral I taken along E thus becomes a function $J(u)$. We set

$$Q(u) = J_{u_h u_k}(0)\, u_h u_k \qquad (h, k = 1, 2, \ldots, \nu)$$

and term $Q(u)$ *the special form associated with* γ taken from $x = a$ to $x = b$.

The following lemma and theorem are established in Morse § 10 — § 13.

Lemma. *The special form $Q(u)$ associated with γ taken from $x = a$ to $x = b$ is given by the formula*

$$Q(u) = \int_a^b (r_{ij}\eta_i'\eta_j' + p_{ij}\eta_i\eta_j)\, dx \qquad (i, j = 1, 2, \ldots, n)$$

where $[\eta(x)]$ gives the y-coordinates along the broken extremal joining the points (8.1) *determined by* (u).

Theorem 3. *If $x = b$ is a conjugate point of $x = a$ of the rth order, the rank of $Q(u)$ is $\nu - r$. If $x = b$ is not conjugate to $x = a$ the rank of $Q(u)$ is ν, and its negative type number is the sum of the orders of the conjugate points of $x = a$ preceding $x = b$.*

§ 9.

The continuous variation of conjugate points and focal points.

By the kth conjugate point of a point $x = a$ on γ is meant the kth conjugate point of $x = a$ following $x = a$, counting conjugate points according to their orders. We shall prove the following lemma.

Lemma. *The kth conjugate point $x = c$ of a point $x = a$ on γ varies continuously with $x = a$ as long as $x = c$ remains on γ.*

Let $x = b$ and $x = b'$ be any two points on γ or on a slight extension of γ, not conjugate to $x = a$, and such that

$$a < b' < c < b.$$

The special form $Q(u)$ of Theorem 3 set up for the segment of γ between $x = a$ and $x = b$ will be non-singular, and of negative type at least k. For a sufficiently small variation of $x = a$, Q will remain non-singular and hence unchanged in type, so that the kth conjugate point will exist and preceed $x = b$.

On the other hand the form $Q(u)$ set up for the segment of γ between $x=a$ and $x=b'$ will be non-singular and of negative type less than k. For a sufficiently small variation of $x=a$, $Q(u)$ will remain non-singular, and hence of type less than k, so that the kth conjugate point of $x=a$ will follow $x=b'$.

The lemma follows from the fact that b and b' may be taken arbitrarily near the given value of c.

We shall now prove the following theorem.

Theorem 4. *The kth conjugate point of a point $x=a$ on γ advances or regresses continuously with $x=a$ as long as it remains on γ.*

According to the preceding lemma the kth conjugate point of $x=a$ varies continuously with $x=a$. To prove that it advances with $x=a$, suppose the theorem false and that when $x=a$ advances to a nearby point $x=a_1$ the kth conjugate point of $x=a$ regresses from $x=c$ to $x=c_1$. Let $x=b$ be a point not conjugate to $x=a$, and between c_1 and c. We are supposing that

$$a < a_1 < c_1 < b < c.$$

The form $Q(u)$ set up for the segment of γ from $x=a$ to $x=b$ will be non-singular, and of negative type less than k, since b precedes c. On the other hand if we choose the point of division x_1 of § 7 as a_1, and set the first n of the variables (u) in $Q(u)$ equal zero, $Q(u)$ will reduce to a non-singular form which will have a negative type number equal to the number of conjugate points of $x=a_1$ preceding $x=b$, or since c_1 precedes b, a negative type number at least k. According to the theory of quadratic forms $Q(u)$ must then be of negative type number at least k. From this contradiction we infer that the kth conjugate point of $x=a$ advances with $x=a$.

Similarly it follows that the kth conjugate point of $x=a$ regresses with $x=a$ and the theorem is proved.

We state the following theorem.

Theorem 5. *The kth focal point of a conjugate family following a point $x=a$ not a focal point, varies continuously with the initial elements g_{ij} at $x=a$, and with the coefficients of the differential equations.*

That the coefficients of the form $H(u)$ vary continuously with the elements g_{ij}, r_{ij} and p_{ij}, follows with the aid of (7.4). One can then repeat the proof of the lemma in this section, referring to focal points where conjugate points are there referred to, and using $H(u)$ in place of $Q(u)$. In this manner one arrives at a proof of the present theorem.

§ 10.

Separation Theorems.

We commence with the following theorem.

Theorem 6. *Let g_{ij} and g_{ij}^0 be respectively the initial elements at $x = a$ of two conjugate families F and F^0. Let P and N be respectively the positive and negative type numbers of the form*

$$D(u) = (g_{ij} - g_{ij}^0) u_i u_j \qquad (i, j = 1, \ldots, n).$$

If q and q^0 are respectively the numbers of focal points of F and F^0 on the interval $a \leqq x \leqq b$, we have

$$q^0 - P \leqq q \leqq q^0 + N. \tag{10.0}$$

It will be sufficient to give the proof for the case that $x = b$ is not a focal point of F or F^0. For if for the given a the theorem is true for every $b > a$ for which $x = b$ is not a focal point, it is clearly true for the special values of b for which $x = b$ is a focal point.

Let the fundamental forms corresponding to F and F^0 taken from $x = a$ to $x = b$, be denoted by $H(u)$ and $H^0(u)$ respectively. According to the lemma of § 7 we have

$$H(u) - H^0(u) = D(u) \qquad (i, j = 1, \ldots, n). \tag{10.1}$$

Now $D(u)$ may be written as a form $D(u)$ of rank P and positive type P, minus a form $N(u)$ of rank N and positive type N. We can therefore write (10.1) in the form

$$H(u) - P(u) = H^0(u) - N(u). \tag{10.2}$$

The form $H^0(u)$ is of negative type q^0, and hence will be negatively definite on a suitably chosen q^0-plane π through the origin in the space of the μ variables (u). Further $P(u)$ will be zero on a suitably chosen $(\mu - P)$-plane π_1 also through the origin. If $q^0 - P > 0$, π will have at least a $(q^0 - P)$-plane π_2 in common with π_1. We see then from (10.2) that $H(u)$ will be negatively definite on π_2. Thus $q \geqq q^0 - P$.

On reversing the roles of H and H^0 one proves that q^0 is at least $q - N$. Hence $q \leqq q^0 + N$. Thus the theorem is proved.

Since $P + N \leqq n$ we have the following corollary.

Corollary 1. *The number of focal points of any conjugate family on a given interval differs from that of any other conjugate family by at most n.*

If we are dealing with a second order differential equation in the plane, any solution of the differential equation not identically zero gives the base of a conjugate family. Here $n = 1$ and the corollary becomes the famous Sturm Separation Theorem.

Corollary 1 applied to that particular conjugate family F which determines the conjugate points of a given point, leads to Corollary 2.

Corollary 2. *If there are q conjugate points of the point $x = a$ on the interval $a < x \leqq b$, there will be at most $q + n$ focal points of any conjugate family on that interval.*

To still further extend Theorem 6 we need the following lemma.

Lemma. *The nullity of the difference form $D(u)$ equals the number of linearly independent extremals common to the two families F and F^0.*

The nullity of the form $D(u)$ is the number of linearly independent solutions $(u_1, \ldots, u_n)$ of the equations

$$(10.3) \qquad D_{u_i} = 2(g_{ij} - g^0_{ij})\, u_j = 0 \qquad (i, j = 1, \ldots, n).$$

If we refer to the definition (5.1) of the initial elements g_{ij}, we see that (10.3) may be written in the form

$$(10.4) \qquad r_{ik}(a)\,[y'_{kj}(a) - y^{0\prime}_{kj}(a)]\, u_j = 0$$

where $y_{ij}(x)$ and $y^0_{ij}(x)$ are respectively elements of bases of F and F^0 which are unitary at $x = a$. But (10.4) is equivalent to the equations

$$(10.5) \qquad y'_{kj}(a)\, u_j = y^{0\prime}_{kj}(a)\, u_j \qquad (k, j = 1, \ldots, n).$$

Since (10.5) gives the conditions that the members of F and F^0 which take on the values $(u_1, \ldots, u_n)$ at $x = a$ be identical with each other, the lemma follows at once.

This leads to the following theorem.

Theorem 7. *If two conjugate families have r linearly independent solutions in common, the number of focal points of the one family on a given interval differs from that of the other family by at most $n - r$.*

According to the lemma the nullity of $D(u)$ is here r, so that in Theorem 6, $P + N = n - r$. The present theorem now follows from (10.0).

To illustrate the general content of Theorem 6 and its corollaries, let us consider an extremal segment g of a regular problem in parametric form in the calculus of variations. (See Morse § 1.) We can transform g into a segment γ of the x axis (Morse § 4), and obtain the following theorem.

Theorem 8. *Suppose g is an extremal in a regular problem in parametric form in the calculus of variations. Let S_1 and S_2 be two regular manifolds of class C''' which cut g transversally at points at which the integrand F is positive. Then the number of focal points of S_1*

on any segment of g differs from the corresponding number for S_2 by at most n.

The proof will be at once clear to the student of the calculus of variations, and will not be given here.

§ 11.

A comparison of coefficients.

Let there be given two self-adjoint differential systems

$$\frac{d}{dx}(r_{ij}y_j') - p_{ij}y_j = 0, \tag{11.1}$$

$$\frac{d}{dx}(\bar{r}_{ij}y_j') - \bar{p}_{ij}y_j = 0 \tag{11.2}$$

satisfying the requirements of § 3. We shall prove the following theorem.

Theorem 9. *Let F and $\bar{F}$ be respectively conjugate families of* (11.1) *and* (11.2). *If F and $\bar{F}$ have the same initial elements g_{ij} at $x = a$, and if*

$$\bar{r}_{ij}\eta_i\eta_j < r_{ij}\eta_i\eta_j \qquad (\eta) \neq 0, \tag{11.3}$$

$$\bar{p}_{ij}\eta_i\eta_j < p_{ij}\eta_i\eta_j \qquad (\eta) \neq 0 \tag{11.4}$$

then the qth focal point of F, if it exists, must be preceded by the qth focal point of $\bar{F}$. Moreover the same result holds if either (11.3) *or* (11.4), *but not both* (11.3) *and* (11.4) *become equalities.*

Let $H(u)$ and $\bar{H}(u)$ be respectively the fundamental forms for F and $\bar{F}$, taken from $x = a$ to $x = b$. Let us suppose $H(u)$ expressed by (7.4). For the same functions $[\eta(x)]$ it follows from the minimizing properties of the integral in (7.4) that

$$\bar{H}(u) \leqq g_{ij}u_iu_j + \int_a^b [\bar{r}_{ij}\eta_i'\eta_j' + \bar{p}_{ij}\eta_i\eta_j]\,dx.$$

From this relation and (7.4), together with (11.3) and (11.4) we see that

$$\bar{H}(u) < H(u) \qquad (u) \neq 0. \tag{11.5}$$

Now holding a fast, place $x = b$ on γ at a focal point of F of the rth order. Suppose b is then preceded by k focal points, counting focal points from $x = a$. It follows from Lemma A of § 7 that $H(u) \leqq 0$ along some $(k + r)$-plane t through the origin. According to (11.5), $\bar{H}(u) < 0$ on t, except at most at the origin.

The negative type number of $\bar{H}(u)$ must then be at least $k + r$, so that the $(k + r)$th focal point of $\bar{F}$ must precede b, and the theorem is proved.

§ 12.

Further comparison of initial conditions.

Theorem 10. *Let F and F^0 be respectively conjugate families of* (11.1) *determined at $x=a$ by initial elements g_{ij} and g^0_{ij}. If*

$$g_{ij}z_iz_j < g^0_{ij}z_iz_j \qquad (i, j = 1, 2, \ldots, n), \quad (z) \neq (0) \tag{12.1}$$

then the kth focal point of F^0 following $x=a$, if it exists, must be preceded by the kth focal point of F.

Let $H(u)$ and $H^0(u)$ be respectively the fundamental forms for F and F^0 taken from $x=a$ to $x=b$. Suppose b is a focal point of F^0. With the aid of (7.4) we see that

$$H(u) - H^0(u) = (g_{ij} - g^0_{ij})\,u_iu_j \quad (i, j = 1, 2, \ldots, n). \tag{12.2}$$

It follows from the final lemma of § 7 that $H^0(u) \leqq 0$ at each point of a q-plane π_q, where q is the sum of the orders of the conjugate points of F^0 on the interval from a to b inclusive, and that at no point on π_q except the origin are the first n of the u_i's zero. Thus on π_q the right member of (12.2) is negative except when $(u) = (0)$. Hence $H(u) < 0$ on π_q except when $(u) = (0)$.

The type number of $H(u)$ must then be at least q, and hence the qth focal point of F must precede $x=b$. The theorem follows at once.

This theorem can be given a geometric form in terms of manifolds S and S_0 which cut the x axis transversally at $x=a$. If we recall that transversality reduces to orthogonality at a point of the x axis, and recall also that

$$g_{ij} = -\bar{x}_{y_iy_j}(0),$$

where $x = \bar{x}(y)$ was our representation of S, we have the following result.

If the difference between the x coordinate of a point on S, and a point on S_0 with the same (y) is a positive definite form in the set (y), then the kth focal point of S_0 following $x=a$, if it exists, must be preceded by the kth focal point of S.

If we have a single second order differential equation

$$\frac{d}{dx}(ry') - py = 0 \qquad (r > 0)$$

the condition (12.1) becomes

$$g_{11} < g^0_{11}. \tag{12.3}$$

If we refer to the definition of the elements g_{ij} we see that (12.3) takes the form

$$y'_{11}(a) < y^{0\prime}_{11}(a)$$

where

$$y_{11}(a) = y^0_{11}(a) = 1.$$

The theorem then becomes a well known comparison theorem.

§ 13.

Principal planes.

In order to obtain a more intimate knowledge of the variation of the focal points with the initial elements of a conjugate family F we introduce a generalization of the principal directions on a surface as follows.

Suppose $x=a$ is not a focal point of F. Let $x=c$ be a focal point of F of the rth order. The curves of F which pass through $x=c$ on γ intersect the n-plane $x=a$ in an r-plane π_r which will be called the *principal r-plane* corresponding to the focal point c.

Suppose now that the initial elements g_{ij} at $x=a$ are functions $\bar{g}_{ij}(\alpha)$ of class C' of a parameter α for α neighboring α_0. For each α we obtain then a conjugate family, say F_α. We need the following lemma.

Lemma. *If as α approaches α_0, certain focal points approach the point $x=c$ as a limit point, the corresponding principal planes on $x=a$ will have all their limit points on the principal plane corresponding to $x=c$.*

Let $y_{ij}(x,\alpha)$ be a base of F_α that is unitary at $x=a$. From the equations (5.1) defining the elements g_{ij} it follows that $y_{ij}(x,\alpha)$ is continuous in both x and α. A curve of F which meets the n-plane $x=a$ in a point $(y)=(d)$ will be given by

$$y_i = d_j y_{ij}(x,\alpha). \tag{13.1}$$

Now the right hand members of (13.1) are continuous in (d), x, and α. The lemma follows readily.

We state now our final comparison theorem.

Theorem 11. *Let $x=c$ be a focal point of the rth order of the family F_{α_0}, and let t be the corresponding principal r-plane on $x=a$. If on t the form*

$$\bar{g}'_{ij}(\alpha_0)\, y_i y_j \qquad (i,j=1,2,\ldots,n) \tag{13.2}$$

is definite, then an increase of α in any sufficiently small neighborhood of α_0 will cause an advance or regression of the r focal points at $x=c$ according as the form (13.2) is positively or negatively definite on t.

Let $H_0(u)$ and $H(u)$ be respectively the fundamental forms corresponding to the families F_{α_0} and F_α, taken on the interval $a \leqq x \leqq c$. From (7.4) we see that if we set (y) equal to the first n variables in the set (u) we have

$$H(u) - H_0(u) = (\alpha - \alpha_0)\, \bar{g}'_{ij}(\bar{\alpha})\, y_i y_j \qquad (i,j=1,\ldots,n) \tag{13.3}$$

where $\bar{\alpha}$ lies between α and α_0.

To be specific suppose the form (13.2) is negatively definite on t. For $(\alpha - \alpha_0)$ positive, and sufficiently small, the right member of (10.3) will still be negatively definite on t.

It will be convenient to call a point (u) *unitary* if $u_i u_i = 1$.

Let us apply Lemma A of § 7 to the family F_{α_0}. We shall prove that $H(u)$ is negatively definite on the q-plane π_q of this lemma, where q is the number of focal points of F_{α_0} on the interval $a \leqq x \leqq c$. To prove this it will suffice to prove that $H(u) < 0$ at all unitary points on π_q.

Now, essentially by hypothesis the right member of (13.3) is negatively definite at the points (y) on t, and accordingly is negative at the unitary points (u) on the r-plane π_r of Lemma A, since the first n coordinates of (u) on π_r give points (y) on t. Hence $H(u) < 0$ at unitary points (u) within a sufficiently small positive distance e of the unitary points on π_r.

At the remaining unitary points on π_q, Lemma A affirms that $H_0(u)$ is negative. Hence at these points $H(u) < 0$, if $\alpha - \alpha_0$ be sufficiently small.

Thus after a sufficiently small increase of α from α_0, $H(u) < 0$ for all unitary points on π_q. Hence $H(u)$ is negatively definite on π_q. Thus the negative type number of $H(u)$ will be at least q. Hence as α increases slightly from α_0, the r focal points originally at $x = c$ regress.

Not only will the focal points at $x = c$ regress when α increases from α_0, but also any small increase of α in the neighborhood of α_0 will correspond to a regression of the focal points in the neighborhood of $x = c$.

This follows from the lemma of this section, because with the aid of that lemma, and with the use of unitary points we see that the form

$$\bar{g}'_{ij}(\alpha)\, y_i y_j \qquad (i, j = 1, \ldots, n)$$

will be negatively definite on each of the principal planes that correspond to focal points near $x = c$, provided $\alpha - \alpha_0$ be sufficiently small. Thus in the proof already given we can replace α_0 by any other value of α, say α_1, sufficiently near α_0, and then show that a small increase of α from α_1 also corresponds to a regression of the focal points near $x = c$.

The case where the form (13.2) is positively definite is treated by first supposing that α decreases, and repeating in essence the preceding proof. The results for an increasing α then follow.

Thus the theorem is proved.

(Eingegangen am 10. 7. 1929.)