APPLICATIONS OF THE THEORY OF QUADRATIC FORMS IN HILBERT SPACE TO THE CALCULUS OF VARIATIONS

Magnus R. Hestenes

1. Introduction. One of the interesting chapters in the calculus of variations is the theory of indices. It has two aspects, the theory in the large and the theory in the small. An important part of the latter is the theory of indices of the second variation, that is, of integrals that are quadratic in their arguments. Such a theory includes the Sturm-Liouville theory for self-adjoint differential systems. The theory of the second variation can be approached from many points of view. It can be developed by means of the theory of differential equations and the associated boundary value problems. A description of this method together with references to the many writers on this subject can be found in the works of Bliss [3; 4] and Reid [14; 15]. The first of these papers by Reid is an excellent introduction to the present paper; in it is found an extensive bibliography on boundary value problems together with a description of various methods of studying such problems. The second variation can also be studied by the use of the theory of broken extremals, as has been done by Morse [13]. A third method is by means of "natural isoperimetric conditions"; this was done by Birkhoff and Hestenes [2] and also by Hazard [7], Karush [10], and Ritcey [17].

The author has been convinced for some time that the theory of the second variation can be obtained from an appropriate theory of quadratic forms in Hilbert space. The purpose of the present paper is to show how this can be done. The theorems in Hilbert space which we shall use can be found in standard references on Hilbert space [18; 19]. In order to apply these results to the calculus of variations, it is more convenient to emphasize the quadratic form, itself, instead of the self-adjoint transformation associated with the quadratic form. For this reason a portion of the paper is devoted to rephrasing known results concerning transformations in terms of quadratic forms.

In order to apply the theory of quadratic forms in Hilbert space to the calculus
of variations, an appropriate generalization of the condition of Legendre is needed.
This condition in the calculus of variations appears in two forms, the weaker and
the stronger form. The generalizations we make are the following: a quadratic form
$Q(x)$ satisfies the Legendre condition in its weaker form if it is lower semicon-
tinuous relative to the weak topology, that is, if $Q(x)$ is expressible in the form
$Q(x) = P(x) - K(x)$, where $P(x)$ is nonnegative and $K(x)$ is continuous relative to
the weak topology. A quadratic form $J(x)$ will be said to satisfy the strengthened
condition of Legendre if it is expressible in the form $J(x) = D(x) - K(x)$, where
$D(x)$ is positive definite and $K(x)$ is continuous relative to the weaker topology.
Essentially this condition is satisfied if and only if weak convergence and con-
vergence of the corresponding values of $J(x)$ imply strong convergence. Forms $J(x)$
which satisfy this latter condition will be called *Legendre forms*. Since we may
take $D(x)^{1/2}$ as a norm, the study of Legendre forms is equivalent to the study of
completely continuous linear transformations. The literature on these transforma-
tions is extensive; see [19]. However, their connection with the calculus of
variations does not appear to have been treated adequately heretofore.

It will be seen in §11 below that Legendre forms have a finite (negative) index
and finite nullity. These indices are fundamental in the calculus of variations.
The nullity, for example, may be used to describe the number of linearly inde-
pendent solutions of a certain differential or integrodifferential equation satisfying
given boundary conditions. The index can be used to describe the number of oscil-
lations of solutions of these differential equations. This latter result is a conse-
quence of the theory of focal points. In §17 below we develop a theory of focal
points which not only yields the standard theory of focal points in the calculus of
variations but also suggests new interpretations for integrodifferential and integral
equations. Of particular interest is the fact that to each resolution of the identity
there corresponds a theory of focal points. When interpreted in terms of differential
systems the theory of focal points yields the Sturm-Liouville theory.

Legendre forms have another property that is of significance in the calculus of
variations. This property will be called *quasi-nonsingularity*. If one examines the
literature in the calculus of variations one finds that this property plays a funda-
mental role in proofs. The form in which this property normally appears is de-
scribed in Theorem 12.3. An interesting by-product of the concept of quasi-non-
singularity is a new proof of the Lagrange multiplier rule given in §14 below.

An attempt has been made to give a sufficient number of illustrations so as to
indicate the connections between theorems and their applications in the calculus

of variations. Only the simplest examples have been used. We have omitted, for example, most of the applications to isoperimetric problems, to problems of Bolza, to parametric problems, to the case of discontinuous solutions, and to multiple integrals. Except in the case of multiple integrals, it is a relatively simple matter to apply the results here given if one is familiar with the problem considered. In the case of multiple integrals further study of the theory in the calculus of variations seems to be necessary to bring the theory to a state of completeness comparable with that for simple integrals.

2. **Hypotheses and preliminary remarks.** Let $\mathcal{A}$ be a linear space over the field of reals. The elements of $\mathcal{A}$, called vectors, will be denoted by $x, y, z, \cdots$. Real numbers, called scalars, will be denoted by $a, b, c, \cdots$. The sum of two vectors x and y will be denoted by $x + y$, and the product of x by a scalar b by bx or xb. A subclass $\mathcal{B}$ of $\mathcal{A}$ that is closed relative to addition and scalar multiplication is called a *linear subclass of* $\mathcal{A}$. By the dimension of $\mathcal{B}$ will be meant the number of linearly independent vectors in $\mathcal{B}$ in a maximal set of such vectors. We shall have occasion to distinguish only between finite and infinite dimensional linear subclasses of $\mathcal{A}$. A set of vectors $x_1, \cdots, x_n$ will be said to *generate* the linear subclass $\mathcal{B}$ of $\mathcal{A}$ comprised of all vectors of the form $a_1 x_1 + \cdots + a_n x_n$. If the vectors $x_1, \cdots, x_n$ are linearly independent, they will be said to form a *basis* for the subclass $\mathcal{B}$ which they generate. A linear subclass $\mathcal{B}$ of $\mathcal{A}$ will be said to be the direct sum of linear subclasses $\mathcal{B}_1, \cdots, \mathcal{B}_n$ if every vector x in $\mathcal{B}$ is expressible uniquely as a sum $x = x_1 + \cdots + x_n$ with x_i in $\mathcal{B}_i$ $(i = 1, \cdots, n)$ and if every such sum is in $\mathcal{B}$.

It will be assumed throughout that we are given a symmetric function (x, y) on $\mathcal{A}\mathcal{A}$ to reals, called the *inner product* of x and y, having the following properties: (a) $(x, x) \geq 0$, the equality holding only in case $x = 0$; (b) $(x, ay + bz) = a(x, y) + b(x, z)$; (c) every Cauchy sequence has a limit; that is, given a sequence $\{x_q\}$ such that

$$\lim_{p = q = \infty} |x_p - x_q| = 0 ,$$

where $|x| = (x, x)^{1/2}$, there is a vector x_0 in $\mathcal{A}$ such that

(2.1) $$\lim_{q = \infty} |x_q - x_0| = 0 .$$

The quantity $|x| \equiv (x, x)^{1/2}$ is called the *norm* or *length* of x and satisfies the relations

$$|x| \geq 0, \quad |ax| = |a| \; |x|, \quad |(x,y)| \leq |x| \; |y|, \quad |x + y| \leq |x| + |y|.$$

The quantity $|x - y|$ denotes the *distance* from x to y.

Two vectors x and y are *orthogonal* if $(x, y) = 0$. A vector x is *orthogonal* to a subclass $\mathcal{B}$ of $\mathcal{Q}$ if it is orthogonal to each vector y in $\mathcal{B}$. Two subclasses $\mathcal{B}$ and $\mathcal{C}$ are said to be orthogonal if every vector x in $\mathcal{B}$ is orthogonal to every vector y in $\mathcal{C}$. The set of all vectors orthogonal to a subclass $\mathcal{B}$ will be called the *orthogonal complement* of $\mathcal{B}$.

A sequence of vectors $\{x_q\}$ will be said to *converge strongly* to a vector x_0, written $x_q \Longrightarrow x_0$, in case (2.1) holds. It will be said to *converge weakly* to x_0, written $x_q \longrightarrow x_0$, in case

(2.2)
$$\lim_{q=\infty} (x_q, y) = (x_0, y)$$

for every vector y in $\mathcal{Q}$. It is *bounded* in case the sequence of its norms $\{|x_q|\}$ is bounded. The symbol $a_q \longrightarrow a_0$ will be used to signify that the sequence of scalars $\{a_q\}$ converges to a_0. By a *closed subset* of $\mathcal{Q}$ will be meant one that is closed relative to strong convergence.

A real-valued function $f(x)$ on $\mathcal{Q}$ is said to be *continuous* if $f(x_q) \longrightarrow f(x_0)$ whenever $x_q \Longrightarrow x_0$. It will be said to be *w-continuous* in case $f(x_q) \longrightarrow f(x_0)$ whenever $x_q \longrightarrow x_0$. If

$$\lim_{q=\infty} \inf f(x_q) \geq f(x_0)$$

whenever $x_q \longrightarrow x_0$, then $f(x)$ is said to be *wls-continuous* on $\mathcal{Q}$, that is, lower semicontinuous on $\mathcal{Q}$ with respect to weak convergence. The function $f(x)$ will be said to be *additive* in case it satisfies the identity

$$f(ax + by) = af(x) + bf(y).$$

A continuous additive function will be said to be *linear* and normally will be denoted by $L(x)$ with or without subscripts.

We shall make frequent use of the standard properties of weak and strong convergence. In particular, one should recall that a closed linear subclass of $\mathcal{Q}$, on which every weakly convergent sequence is strongly convergent, is of finite dimension.

Recall also that if $L_1(x), \cdots, L_k(x)$ are k linear forms, the class $\mathcal{B}$ of all vectors x such that $L_i(x) = 0$ $(i = 1, \cdots, k)$ is a closed linear subclass of $\mathcal{Q}$. Every linear form $L(x)$ vanishing on $\mathcal{B}$ is expressible in the form

$$(2.3) \qquad L(x) = h_1 L_1(x) + \cdots + h_k L_k(x) \ .$$

The multipliers $h_1, \cdots, h_k$ are unique if $L_1(x), \cdots, L_k(x)$ are linearly independent. Given a linear form $L(x)$, there is unique vector y in $\mathcal{G}$ such that $L(x) = (y, x)$.

A function Tx on $\mathcal{G}$ to $\mathcal{G}$ will be called a *linear transformation* if it is continuous and additive in the sense that

$$(2.4) \qquad T(ax + by) = aTx + bTy$$

is an identity. For a linear transformation Tx of this type there is a number M such that

$$(2.5) \qquad |Tx| \leq M|x| \ .$$

If $Tx_q \Rightarrow Tx_0$, whenever $x_q \longrightarrow x_0$, then Tx will be said to be *completely continuous* on $\mathcal{G}$.

A real-valued function $B(x, y)$ on $\mathcal{G}\mathcal{G}$ will be called a *bilinear form* in case it is linear in y for each x and linear in x for each y. To each bilinear form $B(x, y)$ there corresponds a unique pair of linear transformations T and T^*, called *adjoints* of each other, such that

$$(2.6) \qquad B(x, y) = (Tx, y) = (x, T^*y) \ .$$

Observe that $|B(x, y)| \leq M |x| \ |y|$ if M is suitably chosen. Moreover if $x_q \Rightarrow x_0$, and $y_q \longrightarrow y_0$, then $B(x_q, y_q) \longrightarrow B(x_0, y_0)$. If $B(x_q, y_q) \longrightarrow B(x_0, y_0)$ whenever $x_q \longrightarrow x_0$ and $y_q \longrightarrow y_0$, then $B(x, y)$ is said to be *completely continuous* and will normally be denoted by $K(x, y)$. If $K(x, y) = K(y, x)$, then $K(x, y)$ is completely continuous if and only if $K(x) = K(x, x)$ is w-continuous on $\mathcal{G}$, as can be seen by the use of the identity

$$2K(x, y) = K(x + y) - K(x) - K(y) \ .$$

The present paper will be devoted to the study of *quadratic forms* $Q(x) = Q(x, x)$ defined by a symmetric bilinear form $Q(x, y)$. For a quadratic form we have the fundamental identity

$$Q(ax + by) = a^2 Q(x) + 2ab Q(x, y) + b^2 Q(y) \ .$$

A quadratic form is a continuous function of x but is not in general w-continuous. A w-continuous quadratic form will normally be denoted by $K(x)$ and the corresponding bilinear form by $K(x, y)$.

The linear transformation $T(x)$ corresponding to a quadratic form is identical with its adjoint and is accordingly *self-adjoint*. The study of quadratic forms is accordingly equivalent to the study of self-adjoint linear transformations. In the applications to the calculus of variations to be made below, it appears to be more convenient to state our results in terms of quadratic forms rather than in terms of the corresponding linear transformations. The interpretation of these results in terms of transformations will be left to the reader.

3. Example I. In the development of the theory given below the author kept in mind certain standard problems in the calculus of variations which will be described in the next two sections. In the present section we shall consider a space $\mathcal{G}$ that is of interest in the theory of integral equations with a symmetric kernel as well as in the calculus of variations.

The space $\mathcal{G}$ with which we shall be concerned is one in which the vector x is a vector-valued function $x(t) = [x^1(t), \cdots, x^r(t)]$, where $t = (t^1, \cdots, t^p)$ is a point in a p-dimensional Euclidean space restricted to lie in a fixed interval $S: a^\alpha \leq t^\alpha \leq b^\alpha$ ($\alpha = 1, \cdots, p$). Each component $x^j(t)$ is assumed to be a Lebesgue square integrable function on S. The inner product of two vectors x and y is given by the formula

$$(x, y) = \int_S x^j(t) y^j(t) \, dt \qquad\qquad (j = 1, \cdots, r).$$

Here and elsewhere a repeated index in a term denotes summation with respect to that index unless otherwise specified or implied. The norm of x is

$$|x| = \left[\int_S x^j(t) x^j(t) \, dt \right]^{\frac{1}{2}}$$

The class $\mathcal{G}$ together with the inner product (x, y) satisfies the hypotheses made in the last section. Strong convergence is equivalent to convergence in the mean of order two, and weak convergence is weak convergence in the class $\mathcal{L}_2$ of Lebesgue square integrable functions. In fact, $x_q \longrightarrow x_0$ if and only if the integrals

$$\int_S x_q^j(t) x_q^j(t) \, dt \qquad\qquad (q = 1, 2, 3, \cdots)$$

are uniformly bounded and

$$\lim_{q=\infty} \int_{S_0} x_q^j(t)\, dt = \int_{S_0} x_0^j(t)\, dt \qquad (j = 1, \cdots, r)$$

for every subinterval S_0 of S.

We proceed to prove three theorems that will be useful in our applications to the calculus of variations.

THEOREM 3.1. *Let* $A_{jk}(s,t)$ $(j,k = 1, \cdots, r)$ *be* r^2 *Lebesgue square integrable functions on SS. Then the bilinear form*

$$(3.1) \qquad K(x,y) = \int_S \int_S A_{jk}(s,t)\, x^j(s)\, y^k(t)\, ds\, dt$$

is a completely continuous bilinear form on $\mathcal{A}$.

This result is well known. A simple proof can be made as follows: Observe that

$$L(z) = \int_S \int_S A_{jk}(s,t)\, z^{jk}(s,t)\, ds\, dt$$

is a linear form on the class $\mathcal{A}^*$ of Lebesgue square integrable functions $z^{jk}(s,t)$ $(j,k = 1, \cdots, r)$ on SS. By the criterion for weak convergence described above it is seen that the relations $x_q \longrightarrow x_0$, $y_q \longrightarrow y_0$ on $\mathcal{A}$ imply the relation $z_q \longrightarrow z_0$ on $\mathcal{A}^*$, where $z_q^{jk}(s,t) = x_q^j(s)y_q^k(t)$ $(Q = 0,1,2,\cdots)$. Consequently

$$L(z_q) = K(x_q, y_q) \longrightarrow L(z_0) = K(x_0, y_0)\ ,$$

as was to be proved.

THEOREM 3.2. *Let* $R_{jk}(t) = R_{kj}(t)$ $(j,k = 1, \cdots, r)$ *be* $r(r+1)/2$ *essentially bounded integrable functions on* S, *and set*

$$(3.2) \qquad Q(x,y) = \int_S R_{jk}(t)\, x^j(t)\, y^k(t)\, dt\ .$$

Then the quadratic form $Q(x) = Q(x,x)$ *is wls-continuous on* $\mathcal{A}$ *if and only if at almost all points of* S *the inequality*

$$(3.3) \qquad R_{jk}(t)\, a^j a^k \geq 0$$

holds for every set $(a) \neq (0)$. *In fact,* $Q(x)$ *is wls-continuous on* $\mathcal{A}$ *if and only if* $Q(x) \geq 0$ *on* $\mathcal{A}$.

The condition (3.3) is commonly called the *weak Legendre condition* in the

calculus of variations.

Suppose now that $Q(x)$ is *wls*-continuous on $\mathcal{Q}$. Recall that for almost all points t_0 interior to S the relation

$$(3.4) \qquad \lim_{q=\infty} m_q^{-2} \int_{S_q} R_{jk}(t)\, dt = R_{jk}(t_0)$$

holds, where S_q is the set of points in S at a distance at most $1/q$ from t_0, and m_q is the positive square root of the measure of S_q. Let t_0 be an interior point of S at which this limit exists. Consider a set of numbers $a^1, \cdots, a^r$ such that $a^j a^j = 1$. For each integer q let $x_q^j(t) = a^j / m_q$ on S_q and $x_q^j(t) = 0$ elsewhere. The sequence $\{x_q\}$ so defined converges weakly to $x_0 = 0$, and we have $|x_q| = 1$, as one readily verifies. Moreover,

$$Q(x_q) = a^j a^k m_q^{-2} \int_{S_q} R_{jk}(t)\, dt .$$

It follows from (3.4) that

$$\lim_{q=\infty} Q(x_q) = R_{jk}(t_0)\, a^j a^k .$$

Inasmuch as $Q(x)$ is *wls*-continuous we also have

$$\lim_{q=\infty} Q(x_q) \geq Q(x_0) = 0 .$$

The relation (3.3) therefore holds at t_0 and hence almost everywhere on S. Conversely the condition (3.3) implies that $Q(x) \geq 0$ on $\mathcal{Q}$ and hence $Q(x)$ is *wls*-continuous on $\mathcal{Q}$, as we shall see in Lemma 8.1 below.

THEOREM 3.3. *The quadratic form $Q(x)$ defined by (3.2) satisfies the relation*

$$(3.5) \qquad Q(x) \geq h\,|x|^2$$

on $\mathcal{Q}$ with a positive constant h if and only if the inequality

$$(3.6) \qquad R_{jk}(t)\, a^j a^k \geq h\, a^j a^j$$

holds almost everywhere on S.

This result is obtained by applying Theorem 3.2 to the quadratic form

$$Q(x) - h\,|x|^2 = \int_S [R_{jk}(t) - h\delta_{jk}]\, x^j(t)\, x^k(t)\, dt ,$$

where $\delta_{jj} = 1$ and $\delta_{jk} = 0$ $(j \neq k)$. The condition (3.6) is called the *strengthened condition of Legendre*.

Of particular interest is a quadratic form of the type

$$(3.7) \qquad J(x) = K(x) + Q(x) \,,$$

where the bilinear forms $Q(x, y)$ and $K(x, y)$ corresponding to $Q(x)$ and $K(x)$ are given by (3.2) and (3.1) with $A_{jk}(s, t) = A_{kj}(t, s)$. In the case $r = 1$ we have as a special case the quadratic form

$$(3.8) \qquad J(x) = \int_a^b [x(t)]^2 \, dt - \int_a^b \int_a^b A(s, t) x(s) x(t) \, ds \, dt$$

with the symmetric kernel $A(s, t) = A(t, s)$. This quadratic form plays a significant role in the Hilbert-Schmidt theory of integral equations with a symmetric kernel.

4. Example II. In the present paper we shall be primarily interested in the case in which $\mathcal{Q}$ is the totality of arcs x in $(t, x^1, \cdots, x^p)$-space defined by a set of p real-valued functions

$$x : \qquad\qquad x^j(t) \qquad\qquad (a \leq t \leq b, \ j = 1, \cdots, p)$$

that are absolutely continuous and have square integrable derivatives $\dot{x}^j(t)$ on $a \leq t \leq b$. The numbers a and b are held fast. As the inner product of x and y we take

$$(4.1) \qquad (x, y) = x^j(a) y^j(a) + \int_a^b \dot{x}^j(t) \dot{y}^j(t) \, dt \,.$$

The norm $|x| = (x, x)^{1/2}$ of x is accordingly given by the formula

$$(4.2) \qquad |x|^2 = x^j(a) x^j(a) + \int_a^b \dot{x}^j(t) \dot{x}^j(t) \, dt \,.$$

With the help of the remarks made below it is easily verified that the hypotheses made in §2 are satisfied.

LEMMA 4.1. *The relation $x_q \Longrightarrow x_0$ holds if and only if $x_q^j(a) \longrightarrow x_0^j(a)$ and $\dot{x}_q^j(t) \longrightarrow \dot{x}_0^j(t)$ in the mean of order two. Similarly $x_q \longrightarrow x_0$ if and only if $x_q^j(a) \longrightarrow x_0^j(a)$ and $\dot{x}_q^j(t) \longrightarrow \dot{x}_0^j(t)$ weakly in the class of Lebesgue square integrable functions. In either case $x_q^j(t) \longrightarrow x_0^j(t)$ uniformly on $a \leq t \leq b$.*

From this lemma one obtains readily the following:

THEOREM 4.1. *Let* $P_{jk}(t) = P_{kj}(t)$ $(j, k = 1, \cdots, r)$ *be integrable functions and* $Q_{jk}(t)$ *square integrable functions on* $a \leq t \leq b$, *and let*

$$(4.3) \qquad H(x) = A_{jk}x^j(a)x^k(a) + 2B_{jk}x^j(a)x^k(b) + C_{jk}x^j(b)x^k(b) .$$

Then the quadratic form

$$(4.4) \qquad K(x) = H(x) + \int_a^b \left(P_{jk}\, x^j x^k + 2Q_{jk}x^j \dot{x}^k \right) dt$$

is a w-continuous quadratic form on $\mathcal{Q}$.

A second type of w-continuous quadratic form is described in the following:

THEOREM 4.2. *Let* $\Omega(s, t, x, \dot{x}, y, \dot{y})$ *be defined by the formula*

$$\Omega = A_{jk}(s, t)x^j y^k + B_{jk}(s, t)(x^j \dot{y}^k + y^j \dot{x}^k) + C_{jk}(s, t)\dot{x}^j \dot{y}^k ,$$

where $A_{jk}(s, t) = A_{kj}(t, s)$ *are integrable functions of s and t,* $B_{jk}(s, t) = B_{kj}(t, s)$ *are square integrable functions of s and t, and* $C_{jk}(s, t) = C_{kj}(t, s)$ *are essentially bounded integrable functions of s and t. Then the symmetric bilinear form*

$$(4.5) \qquad K(x, y) = \int_a^b \int_a^b \Omega[s, t, x(s), \dot{x}(s), y(t), \dot{y}(t)]\, ds\, dt$$

is completely continuous.

This result follows readily from Lemma 4.1 and Theorem 3.1. In view of Theorem 3.2 we have the following:

THEOREM 4.3. *Let* $R_{jk}(t) = R_{kj}(t)$ $(j, k = 1, \cdots, r)$ *be essentially bounded integrable functions on* $a \leq t \leq b$ *and let* $K(x)$ *be a w-continuous quadratic form on* $\mathcal{Q}$. *Then the quadratic form*

$$(4.6) \qquad J(x) = K(x) + \int_a^b R_{jk}(t)\dot{x}^j(t)\dot{x}^k(t)\, dt$$

is wls-continuous on $\mathcal{Q}$ *if and only if the inequality*

$$(4.7) \qquad\qquad R_{jk}(t)\pi^j \pi^k \geq 0 \qquad\qquad (j, k = 1, \cdots, r)$$

holds almost everywhere on $a \leq t \leq b$ *for every set* $(\pi) \neq (0)$.

Condition (4.7) is known as the Legendre condition.

With the help of Theorem 3.3 and Lemma 8.1 below we obtain:

THEOREM 4.4. *Let $D(x)$ be the quadratic form*

$$(4.8) \qquad D(x) = x^j(a)x^j(a) + \int_a^b R_{jk}(t)\dot{x}^j(t)\dot{x}^k(t)\, dt\,,$$

where $R_{jk}(t) = R_{kj}(t)$ are essentially bounded and integrable on $a \le t \le b$. Then an inequality of the form

$$D(x) \ge h\,|x|^2$$

holds, with a positive constant $h < 1$, if and only if the inequality

$$(4.9) \qquad R_{jk}(t)\,\pi^j\,\pi^k \ge h\pi^j\,\pi^j$$

holds almost everywhere on $a \le t \le b$ for every set $(\pi) \neq (0)$.

Condition (4.9) is known as the strengthened condition of Legendre.

In a later section we shall consider extensions of Theorems 4.3 and 4.4 to the case when our arcs are required to satisfy differential side conditions.

The linear forms on $\mathfrak{A}$ with which we shall be concerned are of the type

$$(4.10) \qquad L(x) = a_k x^k(a) + b_k x^k(b) + \int_a^b [A_k(t)x^k(t) + B_k(t)\dot{x}^k(t)]\, dt\,,$$

where $A_1(t), \cdots, A_r(t)$ are integrable functions and $B_1(t), \cdots, B_r(t)$ are square integrable functions on $a \le t \le b$. Concerning such a form we have:

THEOREM 4.5. *The linear form (4.10) is expressible uniquely in the form $L(x) = (y, x)$, where y is the arc in $\mathfrak{A}$ determined by the conditions*

$$(4.11) \qquad y^k(a) = a_k + b_k + \int_a^b A_k(s)\, ds\,,$$

$$\dot{y}^k(t) = B_k(t) + \int_t^b A_k(s)\, ds + b_k\,.$$

This result is readily verified by substitution.

THEOREM 4.6. *The linear form (4.10) is identically zero on the class $\mathfrak{B}$ of arcs in $\mathfrak{A}$ having $x^j(a) = x^j(b) = 0$ if and only if there exist constants c_k such that the equations*

$$(4.12) \qquad B_k(t) = \int_a^t A_k(s)\, ds + c_k \qquad\qquad (k = 1, \cdots, r)$$

hold almost everywhere on $a \leq t \leq b$.

This result can be established with the help of Theorem 4.5. It is also an easy consequence of the fundamental lemma in the calculus of variations [4, pp. 10-11].

In our illustrations we shall not consider the most general quadratic form that can be constructed but shall limit ourselves to the one normally studied in the calculus of variations. This quadratic form is of the type (4.6) with $K(x)$ defined by (4.4). It is frequently designated by the symbol

$$(4.13) \qquad J(x) = 2q[x(a),\ x(b)] + \int_a^b 2\omega(t, x, \dot{x})\ dt\ ,$$

where $2q$ is the right member of (4.3) and

$$(4.14) \qquad 2\omega = P_{jk}\, x^j x^k + 2Q_{jk} x^j \dot{x}^k + R_{jk} \dot{x}^j \dot{x}^k\ .$$

The corresponding bilinear form is

$$(4.15) \qquad J(x, y) = q_{ka}(x) y^k(a) + q_{kb}(x) y^k(b) + \int_a^b (\omega_{x^k} y^k + \omega_{\dot{x}^k} \dot{y}^k)\ dt\ ,$$

where $q_{ka}(x)$, $q_{kb}(x)$ are the derivatives of $2q[x(a), x(b)]$ with respect to $x^k(a)$, $x^k(b)$, respectively. As was remarked in §2, there is a self-adjoint linear transformation $y = Tx$ such that $J(x, z) = (y, z)$ for all z on $\mathcal{C}$. This transformation is described in the following:

THEOREM 4.7. *Given an arc* x *in* $\mathcal{C}$, *the bilinear form* (4.15) *is expressible uniquely in the form* $J(x, z) = (y, z)$, *where* y *is the arc in* $\mathcal{C}$ *having*

$$.16) \qquad y^k(a) = q_{ka}(x) + q_{kb}(x) + \int_a^b \omega_{x^k}[s, x(s), \dot{x}(s)]\ ds\ ,$$

$$\dot{y}^k(t) = \omega_{\dot{x}^k}[t, x(t), \dot{x}(t)] + \int_t^b \omega_{x^k}[s, x(s), \dot{x}(s)]\ ds + g_{kb}(x)\ .$$

This result follows from Theorem 4.5 with $L(z) = J(x, z)$.

5. Example III. The example with which we shall be concerned is a special e of the general multiple integral problem. Let S denote the interval

$$a^j \leq t^j \leq b^j \qquad\qquad\qquad (j = 1, \cdots,\ r)$$

$t^1, \cdots, t^r$)-space, and denote by $\mathcal{C}$ the class of real-valued functions

$$x : \qquad x(t) = x(t^1, \cdots, t^r) \qquad (t \text{ on } S)$$

having the following properties:

(a) *The functions $x(t)$ are absolutely continuous in each component t^k for almost all $(t^1, \cdots, t^{k-1}, t^{k+1}, \cdots, t^r)$.*

(b) *The functions $x(t)$ vanish on the boundary of S. We can suppose $x(t) = 0$ on the complement of S.*

(c) *The derivatives $\dot{x}_k(t)$ with respect to t^k [and hence also $x(t)$] are square integrable on S.*

Two functions are identified if they differ at most on a set of measure zero on S.

The inner product of two functions x and y in $\mathcal{G}$ is taken to be

$$(x, y) = \int_S \dot{x}_k(t) \, \dot{y}_k(t) \, dt \, .$$

From the results given by Calkin [5], Morrey [11,12], and Hestenes [8], it is seen that the space $\mathcal{G}$ with (x, y) as the inner product defines a Hilbert space.

Let $P(t)$, $Q^k(t)$, $R^{jk}(t) = R^{kj}(t)$ $(j, k = 1, \cdots, p)$ be continuous functions of t on S. In fact we can suppose that they are continuous for all values of t. We shall be interested in the properties of the quadratic form

(5.1) $$J(x) = \int_S \left(P x^2 + 2 Q^k x \dot{x}_k + R^{jk} \dot{x}_j \dot{x}_k \right) dt \, .$$

The following theorems are of interest.

THEOREM 5.1. *The quadratic form*

(5.2) $$K(x) = \int_S \left(P x^2 + 2 Q^k x \dot{x}^k \right) dt$$

is w-continuous on $\mathcal{G}$.

A result of this type under weaker hypothesis has been given by the author [8].

THEOREM 5.2. *The quadratic form*

(5.3) $$D(x) = \int_S R^{jk} \dot{x}_j \dot{x}_k \, dt$$

is wls-continuous on $\mathfrak{A}$ *if and only if the Legendre condition*

$$(5.4) \qquad\qquad R^{jk}(t)\,\pi_j\,\pi_k \geq 0$$

holds on S for all $(\pi) \neq (0)$.

The proof of this result follows from the proof of the necessary condition of Weierstrass for multiple integral problems [6] together with Theorem 7.1 below. The condition of Legendre is a well-known consequence of that of Weierstrass.

THEOREM 5.3. *The quadratic form D(x) defined by* (5.3) *satisfies an inequality of the form*

$$(5.5) \qquad\qquad D(x) \geq h\,|x|^2 \qquad\qquad (h > 0)$$

on $\mathfrak{A}$ *if and only if the strengthened condition of Legendre*

$$(5.6) \qquad\qquad R^{jk}(t)\,\pi_j\,\pi_k > 0$$

holds on S for every set $(\pi) \neq (0)$.

Inasmuch as the coefficients in (5.6) are continuous in t, the inequality (5.6) is equivalent to an inequality of the form

$$(5.7) \qquad\qquad R^{jk}(t)\,\pi_j\,\pi_k \geq h\,\pi_j\,\pi_j$$

holding on S, where h is a suitably chosen positive number. In this event we have

$$(5.8) \qquad\qquad \int_S (R^{jk} - h\,\delta^{jk})\,\dot{x}_j\,\dot{x}_k\,dt \geq 0\,,$$

where $\delta^{kk} = 1$, $\delta^{jk} = 0$ $(j \neq k)$. Consequently (5.5) holds on $\mathfrak{A}$. Conversely if (5.5) holds, the left member of (5.8) is a nonnegative quadratic form and hence is *wls*-continuous (see Lemma 8.1 below). It follows from Theorem 5.2 that (5.7) and hence (5.6) hold as stated.

6. Elementary properties of quadratic forms. The results in this section hold even if the space $\mathfrak{A}$ is not complete. A quadratic form $Q(x)$ will be said to be nonnegative on a subclass $\mathfrak{B}$ of $\mathfrak{A}$ in case the inequality $Q(x) \geq 0$ holds on $\mathfrak{B}$. If $Q(x) > 0$ for every $x \neq 0$ in $\mathfrak{B}$, then $Q(x)$ will be said to be *positive* on $\mathfrak{B}$. The terms "nonpositive" and "negative" are defined similarly by reversing the inequalities.

Two vectors x and y will be said to be *Q-orthogonal* in case $Q(x,y) = 0$. If x

QUADRATIC FORMS IN HILBERT SPACE 539

is Q-orthogonal to every vector y in a subclass $\mathcal{B}$ of $\mathcal{A}$, then x will be said to be Q-*orthogonal* to $\mathcal{B}$. The set of all vectors Q-orthogonal to $\mathcal{B}$ will be called the Q-*orthogonal complement* of $\mathcal{B}$. Two classes $\mathcal{B}$ and $\mathcal{C}$ are said to be Q-*orthogonal* in case each vector in $\mathcal{B}$ is Q-orthogonal to $\mathcal{C}$.

Let $\mathcal{B}$ be a linear subclass of $\mathcal{A}$. Introducing a terminology that is very descriptive in the calculus of variations we shall say that a vector x is a Q-*transversal* of $\mathcal{B}$ if it is in $\mathcal{B}$ and is Q-orthogonal to $\mathcal{B}$. Normally, the symbol $\mathcal{B}_0$ will be used to denote the set of Q-transversals of $\mathcal{B}$.

LEMMA 6.1. *The set of Q-transversals of a linear subclass $\mathcal{B}$ of $\mathcal{A}$ forms a linear subclass $\mathcal{B}_0$ of $\mathcal{B}$. It is closed if $\mathcal{B}$ is closed. Moreover, $Q(x) = 0$ on $\mathcal{B}_0$.*

This result follows readily from the definition of Q-transversals.

LEMMA 6.2. *If $Q(x)$ is nonnegative on a linear subclass $\mathcal{B}$ of $\mathcal{A}$, then a vector x in $\mathcal{B}$ is a Q-transversal of $\mathcal{B}$ if and only if $Q(x) = 0$.*

For if x is in $\mathcal{B}$ and $Q(x) = 0$, then given a vector y in $\mathcal{B}$ one has

$$Q(x + ty) = 2tQ(x, y) + t^2 Q(y) \geq 0$$

identically in t. This is possible only in case $Q(x, y) = 0$, that is, only in case x is a Q-transversal of $\mathcal{B}$. The converse follows from Lemma 6.1.

As a further result we have:

LEMMA 6.3. *Let $\mathcal{C}$ be the set of all the vectors x in $\mathcal{A}$ satisfying a set of m equations*

$$(6.1) \qquad\qquad L_\alpha(x) = 0 \qquad\qquad (\alpha = 1, \cdots, m)$$

defined by linear forms $L_\alpha(x)$. If y is Q-orthogonal to $\mathcal{C}$, then there is a set of multipliers $h_1, \cdots, h_m$ such that we have

$$(6.2) \qquad\qquad Q(y, x) + h_\alpha L_\alpha(x) = 0$$

for all x in $\mathcal{A}$. If $L_1(x), \cdots, L_m(x)$ are linearly independent on $\mathcal{A}$, these multipliers are unique.

This result follows from equation (2.3) for $L(x) = -Q(y, x)$.

LEMMA 6.4. *Let $\mathcal{B}$ be a linear subclass of $\mathcal{A}$ of finite dimension and let $\mathcal{C}$ be its Q-orthogonal complement. Let $\mathcal{A}_0$, $\mathcal{B}_0$, $\mathcal{C}_0$ be the sets of Q-transversals of $\mathcal{A}$,*

$\mathcal{B}$, $\mathcal{C}$ *respectively. Then* :

(a) *A vector* z *is in* $\mathcal{C}_0$ *if and only if it is the sum* $z = x + y$ *of a vector* x *in* $\mathcal{Q}_0$ *and* y *in* $\mathcal{B}_0$.

(b) *If a linear form* $L(x)$ *vanishes on* $\mathcal{B}_0$, *there is a unique vector* y *in* $\mathcal{B}$ *orthogonal to* $\mathcal{B}_0$, *such that* $L(x) = Q(y, x)$ *on* $\mathcal{B}$.

(c) *To every vector* x *in* $\mathcal{Q}$ Q-*orthogonal to* $\mathcal{B}_0$, *there corresponds a unique vector* y *in* $\mathcal{B}$ *orthogonal to* $\mathcal{B}_0$ *such that* $z = x - y$ *is in* $\mathcal{C}$.

(d) *Let* $\mathcal{B}^*$ *be a linear subclass of* $\mathcal{Q}$ *such that every vector common to* $\mathcal{B}_0$ *and* $\mathcal{B}^*$ *is* Q-*orthogonal to* $\mathcal{B}^*$. *If the dimension of* $\mathcal{B}^*$ *exceeds that of* $\mathcal{B}$, *then there is a vector* $x \neq 0$ *in* $\mathcal{B}^*$ *that is* Q-*orthogonal to* $\mathcal{B}$ *and is not in* $\mathcal{B}$.

In order to prove Statement (b), select vectors $y_1, \cdots, y_m$ to be a maximal set of linearly independent vectors in $\mathcal{B}$ orthogonal to $\mathcal{B}_0$. A vector x in $\mathcal{B}$ is in $\mathcal{B}_0$ if and only if $Q(y_\alpha, x) = 0$ ($\alpha = 1, \cdots, m$). Hence if $L(x) \equiv 0$ on $\mathcal{B}_0$, there exists a unique set of multipliers b_α such that for each x in $\mathcal{B}$ we have

$$L(x) = b_\alpha Q(y_\alpha, x) = Q(y, x) ,$$

where $y = y_\alpha b_\alpha$, as was to be proved. Statement (c) follows from (b) by taking $L(\bar{x}) = Q(x, \bar{x})$ and selecting y as described in (b). For in this case $Q(x - y, \bar{x}) = 0$ for all $\bar{x}$ in $\mathcal{B}$.

In order to prove Statement (d), let $\mathcal{B}_1$ be the set of vectors in $\mathcal{B}_0$ belonging to $\mathcal{B}^*$. Select a maximal set of linearly independent vectors $y_1, \cdots, y_m$ in $\mathcal{B}$ orthogonal to $\mathcal{B}_1$. Since the dimension of $\mathcal{B}^*$ exceeds that of $\mathcal{B}$ we can select linearly independent vectors $x_1, \cdots, x_{m+1}$ in $\mathcal{B}^*$ orthogonal to $\mathcal{B}_1$. Choose $a_1, \cdots, a_{m+1}$, not all zero, such that

$$Q(y_\beta, x_\alpha) \, a_\alpha = 0 \qquad (\alpha = 1, \cdots, m + 1; \quad \beta = 1, \cdots, m) .$$

The vector $x = x_\alpha a_\alpha$ has the property described in (d).

In order to prove Statement (a), let $\mathcal{C}^*$ be the Q-orthogonal complement of $\mathcal{B}_0$. By (c) the vectors x in $\mathcal{C}^*$ are of the form $x = y + z$ with y in $\mathcal{B}$ and z in $\mathcal{C}$. It follows that the class $\mathcal{C}_0$ is the class of Q-transversals of $\mathcal{C}^*$ as well as of $\mathcal{C}$. Moreover $\mathcal{B}_0$, $\mathcal{Q}_0$ are subsets of $\mathcal{C}_0$. The orthogonal complement $\mathcal{D}$ of $\mathcal{C}^*$ has no element x, with $x \neq 0$, Q-orthogonal to $\mathcal{B}_0$. By (d) the dimension of $\mathcal{D}$ therefore does not exceed that of $\mathcal{B}_0$. If $x_1, \cdots, x_r$ forms a basis for $\mathcal{D}$ we can find vectors $y_1, \cdots, y_r$ in $\mathcal{B}_0$ such that

$$|Q(x_\alpha, y_\beta)| \neq 0 \qquad\qquad (\alpha, \beta = 1, \cdots, r) .$$

Consider a vector z in C_0. Choose constants b_β such that

$$Q(x_\alpha, y_\beta) \; b_\beta = Q(x_\alpha, z) \;.$$

The vector $x = z - y_\beta b_\beta$ is in C_0, is Q-orthogonal to $x_1, \cdots, x_r$ and to C, and hence also to $\mathcal{A}$. It follows that x is in $\mathcal{A}_0$ and $y_\beta b_\beta$ is in $\mathcal{B}_0$. This proves Statement (a) and thus completes the proof of the lemma.

ILLUSTRATION 1. Turning to Example I in §3 with $J(x)$ defined by (3.8), we see that a vector x is a J-transversal of $\mathcal{A}$ if and only if the equation

$$(6.3) \qquad\qquad x(t) = \int_a^b A(t, s) x(s)\, ds$$

holds almost everywhere on $a \leq t \leq b$.

ILLUSTRATION 2. Turning to Example II in §4 with $J(x)$ defined by (4.13), we see, by Theorem 4.6, that an arc x in $\mathcal{A}$ is J-orthogonal to the class $\mathcal{B}$ of arcs that vanish at $t = a$ and $t = b$ if and only if there exist constants c_k such that

$$(6.4) \qquad\qquad \omega_{\dot{x}^k} = \int_a^t \omega_{x^k}\, dt + c_k \qquad\qquad (k = 1, \cdots, p)$$

almost everywhere on $a \leq t \leq b$. These equations are the Euler equations in integral form; their solutions are called *extremals*. When the strengthened condition of Legendre (4.9) holds, it follows from the theory of differential equations that the extremals form a linear subclass of $\mathcal{A}$ of dimensions $2p$. Given an extremal x we denote the right member of (6.4) by $\xi_k(t)$, thus:

$$(6.5) \qquad\qquad \xi_k(t) = \int_a^t \omega_{x^k}\, dt + c_k \;.$$

This notation will be found useful in further applications.

ILLUSTRATION 3. In Example II of §4 let $\mathcal{B}$ denote the class of arcs x in $\mathcal{A}$ satisfying a set of m linear equations

$$(6.6) \qquad L_\alpha(x) = a_{\alpha k} x^k(a) + b_{\alpha k} x^k(b) = 0 \qquad\qquad (\alpha = 1, \cdots, m) \;.$$

Let $J(x)$ be given by (4.13). An arc x in $\mathcal{B}$ is a J-transversal of $\mathcal{B}$ if and only if it satisfies (6.4) with constants c_k and in addition it satisfies the *transversality conditions,*

(6.7) $-\xi_k(a) + q_{ka} + h_\alpha a_{\alpha k} = 0$,

                               $\xi_k(b) + q_{kb} + h_\alpha b_{\alpha k} = 0$,

where $\xi_k(t)$ is given by equation (6.5), and q_{ka}, q_{kb} are the partial derivatives of $q[x(a), x(b)]$ with respect to $x^k(a)$, $x^k(b)$. For by virtue of Lemma 6.3 there exist constants h_α such that the equation

$$J(x, y) + h_\alpha L_\alpha(y) = 0$$

holds for all y in $\mathcal{G}$. In view of the result in Illustration 2, equations (6.4) hold almost everywhere on $a \leq t \leq b$. Consequently we have

$$J(x, y) = q_{ka} y^k(a) + q_{kb} y^k(b) + \int_a^b \{\omega_{x^k} y^k + \omega_{\dot{x}^k} \dot{y}^k\} dt$$

$$= [q_{ka} - \xi_k(a)] y^k(a) + [q_{kb} + \xi_k(b)] y^k(b) ,$$

where $\xi_k(t)$ is given by (6.5). Using the last two equations one obtains (6.7). The multipliers h_α are unique whenever the matrix $\|a_{\alpha k} \; b_{\alpha k}\|$ has rank m.

The result just described suggested the terminology "J-transversal of $\mathcal{B}$" for an arc in $\mathcal{B}$ that is J-orthogonal to $\mathcal{B}$.

ILLUSTRATION 4. In Example II of §4 denote by $\mathcal{B}$ the set of arcs in $\mathcal{G}$ vanishing at $t = a$ and $t = b$ which also satisfy the isoperimetric conditions

$$L_\alpha(x) = \int_a^b \{A_{\alpha k}(t) x^k(t) + B_{\alpha k}(t) \dot{x}^k(t)\} dt = 0 \qquad (\alpha = 1, \cdots, m) ,$$

in which $A_{\alpha k}$ is integrable and $B_{\alpha k}(t)$ is square integrable on $a \leq t \leq b$. By the use of Theorems 6.3 and 4.6 it is seen that an arc x is J-orthogonal to $\mathcal{B}$ if and only if there exist constants c_k and multipliers h_α such that the equations

(6.8) $\Omega_{\dot{x}^k} = \int_a^t \Omega_{x^k} dt + c_k$

hold almost everywhere on $a \leq t \leq b$, where

$$\Omega = \omega + h_\alpha (A_{\alpha k} x^k + B_{\alpha k} \dot{x}^k) .$$

The multipliers h_α are unique if the linear forms $L_1(x), \cdots, L_m(x)$ are linearly independent on $\mathcal{G}$, that is, if there is no set of multipliers ρ_α and constants d_k such that the equations

$$\rho_u B_{\alpha\,k} = \int_a^t \rho_u A_{\alpha\,k}\,dt + d_k$$

hold almost everywhere on $a \le t \le b$. This is the so-called "normal case" in the calculus of variations.

ILLUSTRATION 5. Turning to Example III in §5, we observe that a function x of class C'' in $\mathcal{C}$ is a J-transversal of $\mathcal{C}$, where $J(x)$ is given by (5.1), if and only if we have

$$\frac{\partial}{\partial t^j}\left(R^{jk}\dot{x}_k + Q^j x\right) = Q^k \dot{x}_k + Px\,.$$

This result follows from the fundamental lemma for multiple integral problems in the calculus of variations.

7. A fundamental property of quadratic forms. A significant portion of the results to be found below is based on the following:

THEOREM 7.1. *Given a quadratic form $Q(x)$ on $\mathcal{C}$, the class $\mathcal{C}$ is expressible in a unique manner as the direct sum of three linear subclasses $\mathcal{C}_-$, $\mathcal{C}_0$, $\mathcal{C}_+$ having the following properties:* (a) *The classes $\mathcal{C}_-$, $\mathcal{C}_0$, $\mathcal{C}_+$ are mutually orthogonal and Q-orthogonal;* (b) $Q(x)$ *is negative on $\mathcal{C}_-$, zero on $\mathcal{C}_0$, and positive on $\mathcal{C}_+$.*

The class $\mathcal{C}_0$ is the class of Q-transversals of $\mathcal{C}$.

In order to prove this result recall that there is a self-adjoint linear transformation T such that $Q(x) = (Tx, x)$. As has been shown by Nagy [19, p. 23] and others, the transformation T is expressible uniquely as the sum $T = T^+ - T^-$ of two self-adjoint linear transformations T^+ and T^- such that the relations

(7.1) $$P(x) = (T^+x, x) \ge 0, \quad N(x) = (T^-x, x) \ge 0\,,$$

(7.2) $$(T^-x, T^+y) = 0$$

hold for all x and y in $\mathcal{C}$. Observe that

(7.3) $$Q(x) = P(x) - N(x), \quad Q(x, y) = P(x, y) - N(x, y)\,.$$

Let $\mathcal{P}$ be the set of P-transversals of $\mathcal{C}$, that is, the set of all vectors x such that $T^+x = 0$. Denote the orthogonal complement of $\mathcal{P}$ by $\mathcal{C}_+$. Similarly let $\mathcal{C}_-$ be the orthogonal complement of the class $\mathcal{N}$ of N-transversals of $\mathcal{C}$. In view of (7.2) we have $T^+T^-x = 0$ and $T^-T^+x = 0$ for all x in $\mathcal{C}$. Consequently T^-x is in $\mathcal{P}$ and T^+x

is in $\mathfrak{n}$. It follows that if x is in $\mathfrak{A}_+$ we have

$$N(x) = (T^- x, x) = 0, \qquad Q(x) = P(x) > 0$$

unless $x = 0$. Similarly

$$P(x) = 0, \qquad Q(x) = -N(x) < 0$$

for all $x \neq 0$ on $\mathfrak{A}_-$. Since $N(x) = 0$ on $\mathfrak{A}_+$ and $N(x) \geq 0$ on $\mathfrak{A}$ it follows from Lemma 6.2 that $\mathfrak{A}_+$ is a subclass of $\mathfrak{n}$ and is accordingly orthogonal to $\mathfrak{A}_-$. Moreover if x is in $\mathfrak{A}_+$ and y is in $\mathfrak{A}_-$, then $T^- x = 0$ and $T^+ y = 0$. Consequently

$$Q(x, y) = (x, T^+ y) - (T^- x, y) = 0 .$$

The classes $\mathfrak{A}_+$ and $\mathfrak{A}_-$ are therefore Q-orthogonal. Let $\mathfrak{A}_0$ be the class of vectors common to $\mathfrak{p}$ and $\mathfrak{n}$. If x is in $\mathfrak{A}_0$, then $T^+ x = T^- x = 0$ and hence $Tx = 0$; that is, x is a Q-transversal of $\mathfrak{A}$. The class $\mathfrak{A}_0$ is therefore Q-orthogonal to $\mathfrak{A}_+$ and $\mathfrak{A}_-$, as well as being orthogonal to these classes. Since $\mathfrak{p}$ is the direct sum of $\mathfrak{A}_0$ and $\mathfrak{A}_-$, and $\mathfrak{A}$ is the direct sum of $\mathfrak{A}_+$ and $\mathfrak{p}$, it follows that $\mathfrak{A}$ is the direct sum of $\mathfrak{A}_+$, $\mathfrak{A}_0$, and $\mathfrak{A}_-$, as was to be proved.

The result just established can be restated as follows:

THEOREM 7.2. *A quadratic form* $Q(x)$ *on* $\mathfrak{A}$ *is expressible uniquely as the difference*

$$Q(x) = P(x) - N(x)$$

of two quadratic forms $P(x)$ *and* $N(x)$ *with the property that* (a) $P(x) = 0$ *on the orthogonal complement of the* N-*transversals of* $\mathfrak{A}$; (b) $N(x) = 0$ *on the orthogonal complement of the* P-*transversals of* $\mathfrak{A}$; (c) $P(x) = N(x) = 0$ *on the class of* Q-*transversals of* $\mathfrak{A}$.

8. Wls-continuous quadratic forms. The quadratic forms in which we shall be interested are wls-continuous. Combining the results given in the present section with those given in Theorems 3.2, 4.3, and 5.2, we see that, in the applications referred to above, a quadratic form is wls-continuous if and only if it satisfies the condition of Legendre in its weaker form.

As a first result we have:

LEMMA 8.1. *If* $Q(x)$ *is nonnegative on a closed linear subclass* $\mathfrak{B}$ *of* $\mathfrak{A}$, *then* $Q(x)$ *is* wls-*continuous on* $\mathfrak{B}$.

In view of the identity

$$Q(x_q) - Q(x_0) \equiv 2Q(x_q - x_0, x_0) + Q(x_q - x_0),$$

we see that whenever $x_q \longrightarrow x_0$ on $\mathcal{B}$, then $Q(x_q - x_0, x_0) \longrightarrow 0$ and

$$\lim_{q=\infty} \inf \; [Q(x_q) - Q(x_0)] = \lim_{q=\infty} \inf \; Q(x_q - x_0) \geq 0$$

inasmuch as $Q(x) \geq 0$ on $\mathcal{B}$.

LEMMA 8.2. *If $Q(x)$ is nonpositive and wls-continuous on a closed linear subclass $\mathcal{B}$ of $\mathcal{A}$ then $Q(x)$ is w-continuous on $\mathcal{B}$.*

For in this case $Q(x)$ and $-Q(x)$ are *wls*-continuous on $\mathcal{B}$ by virtue of Lemma 8.1. Hence $Q(x)$ is *w*-continuous on $\mathcal{B}$.

Combining this result with Theorem 7.1 we obtain:

THEOREM 8.1. *A quadratic form $Q(x)$ is wls-continuous on $\mathcal{A}$ if and only if it is w-continuous on the class $\mathcal{A}_-$ related to $Q(x)$, as described in Theorem 7.1. In particular, if $\mathcal{A}_-$ is of finite dimension then $Q(x)$ is wls-continuous.*

As a consequence we have:

THEOREM 8.2. *A quadratic form $Q(x)$ is wls-continuous on $\mathcal{A}$ if and only if it is expressible as the difference*

(8.1) $$Q(x) = P(x) - K(x)$$

of a nonnegative form $P(x)$ and a w-continuous form $K(x)$. In fact, $K(x)$ can be restricted to be nonnegative and to vanish on the class orthogonal to the P-transversals of $\mathcal{A}$.

If $Q(x)$ is *wls*-continuous on $\mathcal{A}$, then the quadratic form $-N(x)$ described in Theorem 7.2 is nonpositive and *wls*-continuous on $\mathcal{A}$ and hence *w*-continuous on $\mathcal{A}$, by Lemma 8.2. Consequently $Q(x)$ is expressible in the form (8.1). The converse is immediate.

COROLLARY. *If $Q(x)$ is wls-continuous on $\mathcal{A}$ and $Q^*(x) \geq Q(x)$ on $\mathcal{A}$, then $Q^*(x)$ is wls-continuous on $\mathcal{A}$.*

For if $P(x)$ and $K(x)$ are related to $Q(x)$ as described in Theorem 8.2, then

$$P^*(x) = Q^*(x) + K(x) \geq Q(x) + K(x) = P(x) \geq 0$$

on $\mathcal{A}$. Hence $Q^*(x)$ is wls-continuous on $\mathcal{A}$.

THEOREM 8.3. *If* $Q(x)$ *is nonnegative on the orthogonal* (*or* Q-*orthogonal*) *complement of a linear subclass* $\mathcal{C}$ *of* $\mathcal{A}$ *of finite dimension, then* $Q(x)$ *is wls-continuous on* $\mathcal{A}$.

For in this case the dimension of the class $\mathcal{A}_-$ described in Theorem 7.1 cannot exceed that of $\mathcal{C}$ and hence must be finite. Consequently, $Q(x)$ is wls-continuous on $\mathcal{A}$, by Theorem 8.1.

9. Quadratic forms of finite index and nullity. Given a quadratic form $Q(x)$, the dimensions of the classes $\mathcal{A}_-$ and $\mathcal{A}_0$ related to $Q(x)$ as described in Theorem 7.1 will be called, respectively, the *index* i and the *nullity* n of $Q(x)$ *on* $\mathcal{A}$. In the present section we shall be concerned with the case where i is finite and the case where $i + n$ is finite. In these cases $Q(x)$ is wls-continuous on $\mathcal{A}$. Consequently in the examples given in §§3, 4, and 5, the condition of Legendre must hold in its weaker form whenever the quadratic form under consideration is of finite index.

The above definition of index and nullity is valid when $\mathcal{A}$ is replaced by a closed linear subclass $\mathcal{B}$. It fails to hold when $\mathcal{B}$ is not closed since the decomposition described in Theorem 7.1 is based upon the completeness of our space. We shall accordingly define the *nullity of* $Q(x)$ *on a linear subclass* $\mathcal{B}$ *of* $\mathcal{A}$ to be the dimension of the class of Q-transversals of $\mathcal{B}$ and the *index of* $Q(x)$ *on* $\mathcal{B}$ to be the dimension of a maximal linear subclass of $\mathcal{B}$ on which $Q(x)$ is negative. The definition of index is unambiguous in view of the following:

LEMMA 9.1. *Let* $\mathcal{B}$ *be a linear subclass of* $\mathcal{A}$ *and let* $\mathcal{B}_0$ *be the class of its* Q-*transversals. Suppose there exists a maximal linear subclass* $\mathcal{C}$ *of* $\mathcal{B}$ *of finite dimension on which* $Q(x)$ *is negative. Then* $Q(x) \geq 0$ *on the class* $\mathcal{D}$ *of vectors* x *in* $\mathcal{B}$ *that are* Q-*orthogonal to* $\mathcal{C}$, *the inequality holding only in case* x *is in* $\mathcal{B}_0$. *If* $\mathcal{C}^*$ *is a maximal linear subclass of* $\mathcal{B}$ *on which* $Q(x) \leq 0$ *and such that* $\mathcal{C}^*$ *has no vector* $x \neq 0$ *in common with* $\mathcal{B}_0$, *then the dimension of* $\mathcal{C}^*$ *is equal to that of* $\mathcal{C}$.

The first conclusion in the lemma follows from our choice of $\mathcal{C}$ as a maximal linear subclass of $\mathcal{B}$ on which $Q(x)$ is negative. In order to prove the second conclusion suppose that the dimension of $\mathcal{C}^*$ exceeds that of $\mathcal{C}$. Then by Lemma 6.4 (with $\mathcal{A} = \mathcal{B}$) it is seen that there is a vector $x \neq 0$ in $\mathcal{C}^*$ that is Q-orthogonal to $\mathcal{C}$ and hence in $\mathcal{D}$. Clearly $Q(x) = 0$. Hence x must be in $\mathcal{B}_0$ by virtue of the first conclusion. The dimension of $\mathcal{C}^*$ therefore does not exceed that of $\mathcal{C}$. If the dimension of $\mathcal{C}$ were greater than that of $\mathcal{C}^*$, then by Lemma 6.4 we could choose a

vector $x \neq 0$ in C, that is Q-orthogonal to C^* and not in C^*. The vectors of the form $bx + y$, where b is any real number and y is in C^*, would generate a class having the properties ascribed to C^*. Hence C^* could not be maximal. The lemma is therefore established.

The following theorem is valid even if G is not assumed to be complete.

THEOREM 9.1. *The index i of $Q(x)$ on G, if finite, is given by one of the following quantities* : (a) *the dimension of a maximal linear subclass B of G on which $Q(x)$ is negative*; (b) *the dimension of a maximal linear subclass C of G on which $Q(x) \leq 0$ and which contains no nonnull Q-transversal of G*; (c) *the least integer k such that $Q(x) \geq 0$ on the Q-orthogonal complement of a linear subclass C of G of dimension k*; (d) *the least integer k such that $Q(x) \geq 0$ on the orthogonal complement of a linear subclass D of G of dimension k*; (e) *the least integer k such that there exist k linear forms $L_1(x), \cdots, L_k(x)$ such that $Q(x) \geq 0$ whenever $L_\alpha(x) = 0$ $(\alpha = 1, \cdots, k)$.*

The criteria given in (a) and (b) are equivalent by Lemma 9.1. Moreover a class C having the property described in (b) also has the property given in (c). The converse is also true. For if there existed a vector z with $Q(z) > 0$ in the class C described in (c), the set C^* of all vectors x in C that are Q-orthogonal to z would have the following property: A vector x that is Q-orthogonal to C^* is expressible in the form $x = y + bz$, where y is Q-orthogonal to z. Consequently

$$Q(x) = Q(y) + b^2 Q(z) \geq 0 ,$$

contrary to our choice of C. The criteria in (b) and (c) are therefore equivalent. It is easily seen that the criteria given in (d) and (e) are equivalent to that given in (c). The class $\mathrm{D} = \mathrm{G}_-$ described in Theorem 7.1 has the properties described in (c) and (d).

COROLLARY. *If $Q(x)$ is nonnegative on the orthogonal (or Q-orthogonal) complement of a linear subclass C of G of finite dimension k, then $Q(x)$ is of index $i \leq k$.*

We have also the following further result.

THEOREM 9.2. *The sum $m = i + n$ of the index i and the nullity n of $Q(x)$ on G, if finite, is given by each of the following quantities* : (a) *the dimension of a maximal linear subclass B of G on which $Q(x) \leq 0$*; (b) *the least integer k such that $Q(x)$ is positive on the orthogonal complement of a linear subclass B of G of*

dimension k; (c) *the least integer* k *such that there exist* k *linear forms* $L_1(x)$, $\cdots$, $L_k(x)$ *such that* $Q(x) > 0$ *whenever* $x \neq 0$ *and* $L_\alpha(x) = 0$ ($\alpha = 1, \cdots, k$).

This result is readily established with the help of Theorem 9.1. The direct sum $\mathcal{B}$ of the classes $\mathcal{A}_-$ and $\mathcal{A}_0$ described in Theorem 7.1 has the properties described in Theorem 9.2.

COROLLARY. *If* $Q(x)$ *is positive on the orthogonal complement of a linear subspace* $\mathcal{C}$ *of* $\mathcal{A}$ *of finite dimension* k, *the sum of the index and nullity of* $Q(x)$ *on* $\mathcal{A}$ *does not exceed* k.

The following result is immediate.

THEOREM 9.3. *If* $\mathcal{A}^*$ *is a linear subclass of* $\mathcal{A}$, *and* i, i^* *are the indices and* n, n^* *the nullities of* $Q(x)$ *on* $\mathcal{A}$, $\mathcal{A}^*$, *respectively, then*

$$(9.1) \qquad\qquad i^* \leq i, \quad i^* + n^* \leq i + d \leq i + n,$$

where d *is the dimension of the class* $\mathcal{D}$ *of vectors that are simultaneously* Q-*transversals of* $\mathcal{A}$ *and* $\mathcal{A}^*$.

An analogous result is the following:

THEOREM 9.4. *Suppose* $Q^*(x) \geq Q(x)$ *on* $\mathcal{A}$. *If* i, i^* *are the indices and* n, n^* *are the nullities of* $Q(x)$, $Q^*(x)$ *on* $\mathcal{A}$, *then* (9.1) *holds, where* d *is the dimension of the class* $\mathcal{D}$ *of the vectors* x *that are simultaneously* Q-*transversals and* Q^*-*transversals of* $\mathcal{A}$. *If* $Q^*(x) > Q(x)$ *for all* $x \neq 0$ *on* $\mathcal{A}$, *then*

$$(9.2) \qquad\qquad\qquad i^* + n^* \leq i .$$

A somewhat more complete set of inequalities than those given in Theorem 9.3 can be obtained when $\mathcal{A}^*$ is the orthogonal complement of a linear subclass of $\mathcal{A}$ of finite dimension. The result is described in the following:

THEOREM 9.5. *Let* $L_1(x), \cdots, L_k(x)$ *be* k *linearly independent linear forms on* $\mathcal{A}$ *and let* $\mathcal{A}^*$ *be the set of all vectors* x *such that*

$$L_\alpha(x) = 0 \qquad\qquad\qquad (\alpha = 1, \cdots, k) .$$

Then the numbers i, i^*, n, n^*, d *described in Theorem 9.3 satisfy, besides* (9.1), *the further relation*

$$(9.3) \qquad\qquad\qquad i + n \leq i^* + d + k .$$

10. Nonsingular and positive definite quadratic forms. A quadratic form $Q(x)$ will be said to be *nonsingular on a linear subclass* $\mathcal{B}$ *of* $\mathcal{A}$ if given a linear form $L(x)$ there is a unique vector y in $\mathcal{B}$ such that the relation

$$(10.1) \qquad\qquad L(x) = Q(y, x)$$

holds for all x in $\mathcal{B}$. It will be said to be *positive definite on* $\mathcal{B}$ if there is a positive number h such that the inequality

$$(10.2) \qquad\qquad Q(x) \geq h \, |x|^2$$

holds on $\mathcal{B}$. Normally, a quadratic form that is positive definite on $\mathcal{A}$ will be denoted by $D(x)$ and the corresponding bilinear form by $D(x, y)$.

LEMMA 10.1. *If $Q(x)$ is nonsingular on a linear subclass $\mathcal{B}$ of $\mathcal{A}$, then $\mathcal{B}$ is closed and $\mathcal{A}$ is the direct sum of $\mathcal{B}$ and its Q-orthogonal complement $\mathcal{C}$.*

For given a vector x_0 in $\mathcal{A}$, there is a unique vector y_0 in $\mathcal{B}$ such that $Q(x_0, y) = Q(y_0, y)$ for all y in $\mathcal{B}$. Consequently $z_0 = x_0 - y_0$ is in $\mathcal{C}$, and $\mathcal{A}$ is the direct sum of $\mathcal{B}$ and $\mathcal{C}$.

To show that $\mathcal{B}$ is closed, consider a sequence $\{y_q\}$ of vectors in $\mathcal{B}$ converging to a vector x_0 in $\mathcal{A}$. Choose y_0 in $\mathcal{B}$ and z_0 in $\mathcal{C}$ such that $x_0 = y_0 + z_0$. Choose $\bar{y}$ in $\mathcal{B}$ such that $(x_0 - y_0, y) = Q(\bar{y}, y)$ for all y in $\mathcal{B}$. Since y_q is in $\mathcal{B}$ we have

$$|x_0 - y_0|^2 = \lim_{q=\infty} (x_0 - y_0, \, y_q - y_0) = \lim_{q=\infty} Q(\bar{y}, \, y_q - y_0)$$

$$= \lim_{q=\infty} Q(\bar{y}, \, y_q - x_0) = 0 \, ,$$

Hence $x_0 = y_0$, as was to be proved.

Criteria for positive definiteness are given in the following:

THEOREM 10.1. *If a positive quadratic form $D(x)$ has one of the following properties on $\mathcal{A}$ it has them all:*

(a) *$D(x)$ is positive definite.*

(b) *$x_q \Rightarrow 0$ whenever $D(x_q) \longrightarrow 0$.*

(c) *$x_q \Rightarrow x_0$ whenever $x_q \longrightarrow x_0$ and $D(x_q) \longrightarrow D(x_0)$.*

(d) *If $\{D(x_q)\}$ is bounded, then $\{x_q\}$ converges weakly in subsequence.*

(e) *If $\{D(x_q, y)\}$ is bounded for each y in $\mathcal{A}$, then $\{x_q\}$ converges weakly in subsequence.*

(f) *$D(x)$ is nonsingular on $\mathcal{A}$.*

(g) *If $\mathcal{B}$ is a closed linear subspace of $\mathcal{A}$, then $\mathcal{A}$ is the direct sum of $\mathcal{B}$ and its D-orthogonal complement.*

Clearly (a) implies (b). Moreover (b) implies (c) by virtue of the identity

$$D(x_q) - D(x_0) = 2D(x_q - x_0, \ x_0) + D(x_q - x_0) .$$

If (c) failed to imply (a) we could select vectors x_q, x_0 such that $|x_q| = 1$, $D(x_q)$ $< 1/q$, $x_q \longrightarrow x_0$ and hence such that

$$0 = \lim_{q=\infty} D(x_q) \geq D(x_0) = 0 .$$

Consequently, $x_0 = 0$, so that $x_q \Longrightarrow 0$ by (c), contrary to the relations $|x_q| = 1$.

If (a) holds and $\{D(x_q)\}$ is bounded, then $\{|x_q|\}$ is bounded. Consequently (a) implies (d). If (d) failed to imply (a) we could select vectors x_q, x_0 such that $1 = D(x_q) \leq (1/q) |x_q|^2$ and such that $x_q \longrightarrow x_0$, which is impossible.

In order to show that (a) or (d) implies (f), observe that the function $f(x) = D(x)$ $- 2L(x)$, where $L(x)$ is linear, is wls-continuous, and has a bounded minimizing sequence and hence attains its minimum at a vector y. Setting $g(t) = f(y + tx)$, we have

$$g'(0) = 2D(y, x) - 2L(x) = 0$$

for all vectors x in $\mathcal{A}$. Moreover, y is unique since D is positive. Hence (f) holds.

If (f) holds, then (z,x) is expressible in the form $(z,x) = D(y,x)$, and conversely. A sequence $\{D(y, x_q)\}$ is therefore bounded for all y in $\mathcal{A}$ if and only if $\{(z, x_q)\}$ is bounded for every z in $\mathcal{A}$ and hence if and only if $\{|x_q|\}$ is bounded. Consequently (e) follows from (f). Since $|D(x, y)|^2 \leq D(x) D(y)$, property (e) implies (d) and hence (a).

Using Lemma 10.1, we see that (g) follows from (f) and (a). Conversely (g) implies (f). For given a linear form $L(x) = (x_0, x)$ with $x_0 \neq 0$, let $\mathcal{B}$ be the class of vectors having $L(x) = 0$. Select y_0 in $\mathcal{B}$, and z_0 in the D-orthogonal complement $\mathcal{C}$ of $\mathcal{B}$, so that $x_0 = y_0 + z_0$. Clearly $z_0 \neq 0$. Since $D(z_0, x) = 0$ whenever $L(x) = 0$, there is a constant k such that $D(z_0, x) = kL(x)$. Consequently, $L(x) = D(z, x)$, where $z = (1/k) z_0$.

LEMMA 10.2. *Let $\mathcal{B}$ and $\mathcal{C}$ be Q-orthogonal subclasses of $\mathcal{A}$ whose direct sum is $\mathcal{A}$. Then $Q(x)$ is nonsingular on $\mathcal{A}$ if and only if $Q(x)$ is nonsingular on $\mathcal{B}$ and on $\mathcal{C}$.*

For consider a linear form $L(x)$. If $Q(x)$ is nonsingular on $\mathcal{A}$, there is a unique vector x_0 in $\mathcal{A}$ such that $L(x) = Q(x_0, x)$ on $\mathcal{A}$. Select y_0 in $\mathcal{B}$ and z_0 in $\mathcal{C}$ such that $x_0 = y_0 + z_0$. Then $L(y) = Q(y_0 + z_0, y) = Q(y_0, y)$ for all y in $\mathcal{B}$. Similarly

$L(z) = Q(z_0, z)$ on $\mathbb{C}$. It is easily seen that these representations are unique. Conversely if $Q(x)$ is nonsingular on $\mathfrak{B}$ and $\mathbb{C}$, we may choose y_0 in $\mathfrak{B}$ and z_0 in $\mathbb{C}$ such that $L(y) = Q(y_0, y)$ on $\mathfrak{B}$ and $L(z) = Q(z_0, z)$ on $\mathbb{C}$. Since $\mathfrak{A}$ is the direct sum of the Q-orthogonal classes $\mathfrak{B}$ and $\mathbb{C}$ we have $L(x) = Q(y_0 + z_0, x)$ on $\mathfrak{A}$. Again this representation is unique, and the lemma is proved.

THEOREM 10.2. *If $Q(x)$ is nonsingular on every closed linear subclass of $\mathfrak{A}$, then either $Q(x)$ or $-Q(x)$ is positive definite on $\mathfrak{A}$.*

Suppose there is a vector $x_0 \neq 0$ having $Q(x_0) = 0$. Then the Q-orthogonal complement $\mathfrak{B}$ of x_0 is closed and contains x_0. If y is in $\mathfrak{B}$, we have $Q(y + x_0, x) = Q(y, x)$ for every x in $\mathfrak{B}$. Consequently $Q(x)$ cannot be nonsingular on $\mathfrak{B}$, contrary to our assumption. It follows that either $Q(x)$ or $-Q(x)$ is positive and nonsingular on $\mathfrak{A}$ and hence positive definite on $\mathfrak{A}$, as was to be proved.

In a similar manner we may prove:

THEOREM 10.3. *If for every closed linear subclass $\mathfrak{B}$ of $\mathfrak{A}$, the class $\mathfrak{A}$ is the direct sum of $\mathfrak{B}$ and its Q-orthogonal complement, then either $Q(x)$ or $-Q(x)$ is positive definite on $\mathfrak{A}$.*

Examples of positive definite forms are described in Theorems 3.3, 4.4, and 5.3.

11. Legendre forms. A quadratic form $J(x)$ will be called a *Legendre form* in case (a) it is *wls*-continuous on $\mathfrak{A}$ and (b) $x_q \Longrightarrow x_0$ whenever $x_q \longrightarrow x_0$ and $J(x_q) \longrightarrow J(x_0)$. A Legendre form normally will be denoted by $J(x)$ and the corresponding bilinear form by $J(x, y)$. It will be seen that, in the applications to the calculus of variations, Legendre forms are those that satisfy the strengthened condition of Legendre.

In view of (c) in Theorem 10.1 we have:

THEOREM 11.1. *A positive Legendre form is positive definite.*

As a further result we have:

THEOREM 11.2. *A linear subspace $\mathfrak{B}$ of $\mathfrak{A}$ on which a Legendre form $J(x)$ is nonpositive is of finite dimension.*

For in this case $J(x)$ is w-continuous on $\mathfrak{B}$, by Lemma 8.2. Consequently, whenever the relation $x_q \longrightarrow x_0$ holds on $\mathfrak{B}$ we also have $J(x_q) \longrightarrow J(x_0)$ and hence also $x_q \Longrightarrow x_0$, by property (b) of $J(x)$. It follows that weak and strong convergence

are equivalent on $\mathfrak{B}$. This is possible only if $\mathfrak{B}$ is of finite dimension, as was to be proved.

In view of this result the classes $\mathfrak{A}_0$ and $\mathfrak{A}_-$ related to a Legendre form $J(x)$, as described in Theorem 7.1, are of finite dimension. Hence we have:

THEOREM 11.3. *A Legendre form is of finite index and nullity.*

Combining Theorems 11.3 and 11.1 we obtain:

THEOREM 11.4. *A quadratic form $J(x)$ is a Legendre form on $\mathfrak{A}$ if and only if there is a linear subset $\mathfrak{B}$ of $\mathfrak{A}$ of finite dimension such that $J(x)$ is positive definite on the orthogonal complement of $\mathfrak{B}$. If $J(x)$ is a Legendre form on the orthogonal complement of a linear subclass $\mathfrak{B}$ of $\mathfrak{A}$ of finite dimension, then $J(x)$ is a Legendre form on $\mathfrak{A}$.*

The following result is immediate.

THEOREM 11.5. *The sum $J(x) + K(x)$ of a Legendre form $J(x)$ and a w-continuous quadratic form $K(x)$ is again a Legendre form.*

Using this fact we can prove:

THEOREM 11.6. *A quadratic form $J(x)$ is a Legendre form on $\mathfrak{A}$ if and only if it is expressible as the difference*

$$(11.1) \qquad\qquad J(x) = D(x) - K(x)$$

of a positive definite form $D(x)$ and a w-continuous form $K(x)$. In fact $K(x)$ can be restricted to be nonnegative on $\mathfrak{A}$.

If $J(x)$ is of the form (11.1) it is a Legendre form, by virtue of Theorem 11.5. In order to prove the converse, let $x_1, \cdots, x_n$ be a basis for the J-orthogonals of $\mathfrak{A}$. Select $P(x)$, $N(x)$ related to $J(x)$, as described in Theorem 7.2. Then $J(x) = P(x) - N(x)$, and

$$D(x) = P(x) + N(x) + (x_\alpha, x)(x_\alpha, x)$$

is positive on $\mathfrak{A}$. As was seen in proof of Theorem 8.2, the form $N(x)$ is w-continuous on $\mathfrak{A}$ and hence so is the form

$$K(x) = D(x) - J(x) = 2N(x) + (x_\alpha, x)(x_\alpha, x).$$

Since $D(x)$ differs from $J(x)$ by a w-continuous form, $D(x)$ is a Legendre form on $\mathfrak{A}$.

Since $D(x)$ is positive, it is positive definite, by Theorem 11.1, and Theorem 11.6 is thus proved.

COROLLARY 1. *If $J(x)$ is a Legendre form, then there is a nonnegative w-continuous form $K(x)$ such that $J(x) > 0$ for all $x \neq 0$ in $\mathbb{G}$ having $K(x) = 0$.*

COROLLARY 2. *If $J(x)$ is a Legendre form on $\mathbb{G}$ and $J^*(x)$ is a quadratic form such that $J^*(x) \geq J(x)$ in $\mathbb{G}$, then $J^*(x)$ is a Legendre form on $\mathbb{G}$.*

For if $D(x), K(x)$ are related to $J(x)$ as described in the theorem, then

$$D^*(x) = J^*(x) + K(x) \geq J(x) + K(x) = D(x).$$

Consequently, $D^*(x)$ is positive definite on $\mathbb{G}$, and $J^*(x)$ is a Legendre form, as was to be proved.

The condition (b) in the definition of Legendre forms characterizes Legendre forms, apart from sign, as is seen from the following:

THEOREM 11.7. *If a quadratic form $Q(x)$ has the property that $x_q \Longrightarrow x_0$ whenever $x_q \longrightarrow x_0$ and $Q(x_q) \longrightarrow Q(x_0)$, then either $Q(x)$ or $-Q(x)$ is a Legendre form on $\mathbb{G}$.*

It is sufficient to show that either $Q(x)$ or $-Q(x)$ is wls-continuous on $\mathbb{G}$. If this were not so we could select sequences $\{y_q\}$, $\{z_q\}$ converging weakly to vectors y_0, z_0 such that $Q(y_q)$, $Q(y_q, z_q)$, $Q(z_q)$ converge respectively to numbers A, B, C, such that

(11.2) $$A < Q(y_0), \quad C > Q(z_0).$$

The equation

(11.3) $$[A - Q(y_0)]a^2 + 2a[B - Q(y_0, z_0)] + C - Q(z_0) = 0$$

would accordingly have two distinct real roots a_1 and a_2. Then we would have

$$x_{q\alpha} = a_\alpha y_q + z_q \longrightarrow x_{0\alpha} = a_\alpha y_0 + z_0 \qquad (\alpha = 1, 2).$$

Moreover, since

$$Q(x_{q\alpha}) = Q(y_q)a_\alpha^2 + 2 a_\alpha Q(y_q, z_q) + Q(z_q),$$

we would have, by (11.3),

$$\lim_{q=\infty} Q(x_{q\alpha}) = A a_\alpha^2 + 2 a_\alpha B + C = Q(x_{0\alpha}).$$

According to our hypothesis, $x_{q\alpha} \Longrightarrow x_{0\alpha}$. Since $a_1 \neq a_2$ this is possible only in case $y_q \Longrightarrow y_0$ and $z_q \Longrightarrow z_0$. But this would imply that $A = Q(y_0)$, $C = Q(z_0)$, contrary to (11.2). This proves the theorem.

ILLUSTRATION 1. Turning to Example II in §4, we see by virtue of Theorems 4.1 and 4.4 that the integral $J(x)$ defined by (4.6) is a Legendre form on $\mathbb{G}$ if and only if the strengthened condition of Legendre (4.9) holds. In fact by virtue of Theorem 11.4 the quadratic form $J(x)$ is a Legendre form on a linear subclass $\mathbb{B}$ of arcs in $\mathbb{G}$ satisfying a set of conditions of the form

$$L_\alpha(x) = a_{\alpha j} x^j(a) + b_{\alpha j} x^j(b) + \int_a^b \{A_{j\alpha}(t) x^j + B_{j\alpha}(t) \dot{x}^j\} dt = 0$$

$$(\alpha = 1, \cdots, m)$$

if and only if the strengthened condition of Legendre (4.9) holds. If the arcs are required to satisfy a system of differential equations, the criterion (4.9) must be modified, as will be seen in §14 below.

ILLUSTRATION 2. Turning to Example III in §5 we see by Theorems 5.1 and 5.3 that the integral $J(x)$ defined by (5.1) is a Legendre form if and only if the strengthened condition (5.6) of Legendre holds on S.

These illustrations serve to justify the nomenclature "Legendre form."

12. Quasi-nonsingular quadratic forms. A quadratic form $Q(x)$ will be said to be *nondegenerate* on a linear subclass $\mathbb{B}$ of $\mathbb{G}$ if there exists no nonnull Q-transversal of $\mathbb{B}$. It will be said to be *quasi-nonsingular* on $\mathbb{G}$ if it is nonsingular on each closed linear subclass on which it is nondegenerate. If $Q(x)$ is quasi-nonsingular on $\mathbb{G}$ it is quasi-nonsingular on every closed linear subset of $\mathbb{G}$.

The condition of quasi-nonsingularity can be restated as described in the following:

THEOREM 12.1. *A quadratic form $Q(x)$ is quasi-nonsingular on $\mathbb{G}$ if and only if given a closed linear subclass $\mathbb{B}$ of $\mathbb{G}$ and a linear form $L(x)$ that vanishes identically on the class $\mathbb{B}_0$ of Q-transversals of $\mathbb{B}$ there is a vector y in $\mathbb{B}$ such that the equation*

$$L(x) = Q(y, x)$$

holds for every vector x in $\mathbb{B}$. The vector y can be chosen to be orthogonal to $\mathbb{B}_0$, and if so chosen is unique.

A second criterion for quasi-nonsingularity is given in the following:

THEOREM 12.2. *A quadratic form $Q(x)$ is quasi-nonsingular on $\mathfrak{A}$ if and only if, given a closed linear subclass $\mathfrak{B}$ of $\mathfrak{A}$ on which $Q(x)$ is nondegenerate, $\mathfrak{A}$ is the direct sum of $\mathfrak{B}$ and its Q-orthogonal complement.*

The necessity of this criterion follows from Theorem 10.1. In order to prove its sufficiency, let $\mathfrak{B}$ be a closed linear subclass of $\mathfrak{A}$ on which $Q(x)$ is nondegenerate, and consider a linear form $L(x)$ that does not vanish identically on $\mathfrak{B}$. Denote by $\mathfrak{C}$ the subset of $\mathfrak{B}$ on which $L(x) = 0$. We shall show first that there is a vector $y_0 \neq 0$ in $\mathfrak{B}$ which is Q-orthogonal to $\mathfrak{C}$. A nonnull Q-transversal of $\mathfrak{C}$ has this property. If $\mathfrak{C}$ possesses no Q-transversal, $\mathfrak{B}$ is the direct sum of $\mathfrak{C}$ and its Q-orthogonal complement relative to $\mathfrak{B}$, by virtue of our hypotheses. Consequently, in this case also there is a vector $y_0 \neq 0$ in $\mathfrak{B}$ which is Q-orthogonal to $\mathfrak{C}$. By virtue of Theorem 6.3 with $\mathfrak{A} = \mathfrak{B}$, there is a constant h such that $Q(y_0, x) = hL(x)$ on $\mathfrak{B}$. Moreover $h \neq 0$ since otherwise y_0 would be a Q-transversal of $\mathfrak{B}$. Setting $y = y_0/h$, we have $L(x) = Q(y, x)$ on $\mathfrak{B}$. Inasmuch as $Q(x)$ is nondegenerate on $\mathfrak{B}$, y is unique and $Q(x)$ is nonsingular on $\mathfrak{B}$, as was to be proved.

The criterion just established can be restated as follows:

THEOREM 12.3. *A quadratic form $Q(x)$ is quasi-nonsingular on $\mathfrak{A}$ if and only if it has the following property: Let $\mathfrak{B}$ be a closed linear subclass of $\mathfrak{A}$ and denote by $\mathfrak{B}_0$ the class of Q-transversals of $\mathfrak{B}$. Every vector x that is Q-orthogonal to $\mathfrak{B}_0$ is expressible in the form $x = y + z$, where y is in $\mathfrak{B}$ and z is Q-orthogonal to $\mathfrak{B}$. The vector y can be chosen to be orthogonal to $\mathfrak{B}_0$, and if so chosen is unique.*

Let $Q(x)$ be quasi-nonsingular on $\mathfrak{A}$, $\mathfrak{B}$ a closed linear subclass of $\mathfrak{A}$, $\mathfrak{B}_0$ the class of its Q-transversals, and $\mathfrak{B}_1$ the orthogonal complement of $\mathfrak{B}_0$ relative to $\mathfrak{B}$. Every vector x in $\mathfrak{A}$ is expressible in the form $x = y + z$, with y in $\mathfrak{B}_1$ and z Q-orthogonal to $\mathfrak{B}_1$. Moreover, if x is Q-orthogonal to $\mathfrak{B}_0$, then z is Q-orthogonal to $\mathfrak{B}_0$ and hence to $\mathfrak{B}$. The criterion in Theorem 12.3 is therefore a necessary condition for quasi-nonsingularity. It is also sufficient, by Theorem 12.2.

THEOREM 12.4. *A nonnegative quadratic form $Q(x)$ on $\mathfrak{A}$ is quasi-nonsingular on $\mathfrak{A}$ if and only if it is positive definite on the class of vectors orthogonal to the Q-transversals of $\mathfrak{A}$.*

As a further result we have:

THEOREM 12.5. *If $Q(x)$ is quasi-nonsingular on the orthogonal complement C of a linear subspace B of finite dimension, then $Q(x)$ is quasi-nonsingular on $\mathcal{A}$.*

For let $\mathcal{A}^*$ be a closed linear subclass of $\mathcal{A}$ on which $Q(x)$ is nondegenerate. We shall show that $Q(x)$ is nonsingular on $\mathcal{A}^*$. To this end let C^* be the subclass of vectors in $\mathcal{A}^*$ orthogonal to B. Then C^* is a subclass of C. Denote by C_0^* the class of Q-transversals of C^*. Since $\mathcal{A}^*$ is nondegenerate and B is of finite dimension, it follows readily that C_0^* is of finite dimension. We may suppose that C_0^* is of dimension zero, since we can replace B by the algebraic closure of B and C_0^*. Since $Q(x)$ is quasi-nonsingular on C it is nonsingular on C^*. Consequently $\mathcal{A}^*$ is the direct sum of C^* and the class B^* of Q-orthogonals of C^* in $\mathcal{A}^*$. The class B^* is of finite dimension. Moreover $Q(x)$ is nondegenerate on B^* since otherwise $Q(x)$ would be degenerate on $\mathcal{A}^*$. By virtue of Lemma 6.4, $Q(x)$ is nonsingular on B^*. The theorem now follows from Lemma 10.2 with $\mathcal{A} = \mathcal{A}^*$.

COROLLARY. *Let $L_1(x), \cdots, L_p(x)$ be linear forms, and let $A_{\alpha\beta} = A_{\beta\alpha}$ $(\alpha, \beta = 1, \cdots, p)$ be a set of $p(p+1)/2$ real numbers. Then the quadratic form*

$$P(x) = A_{\alpha\beta} L_\alpha(x) L_\beta(x) + Q(x)$$

is quasi-nonsingular on $\mathcal{A}$ if and only if $Q(x)$ is quasi-nonsingular on $\mathcal{A}$.

Combining Theorems 12.5, 11.4, 10.1, and 9.2, we obtain:

THEOREM 12.6. *A quadratic form $J(x)$ is a Legendre form on $\mathcal{A}$ if and only if it is quasi-nonsingular and of finite index and nullity on $\mathcal{A}$.*

This result can be stated in another form:

THEOREM 12.7. *A quadratic form $J(x)$ is a Legendre form on $\mathcal{A}$ if and only if it is quasi-nonsingular, wls-continuous and of finite nullity on $\mathcal{A}$.*

As we shall see presently the Lagrange multiplier rule is a consequence of the following:

THEOREM 12.8. *Let $Q(x)$ be a quasi-nonsingular quadratic form on $\mathcal{A}$, and let B be the set of Q-transversals of $\mathcal{A}$. If $P(x)$ is a quadratic form on $\mathcal{A}$ and x is a vector that is P-orthogonal to B, then there is a vector y in $\mathcal{A}$ such that the relation*

$$P(x, z) = Q(y, z)$$

holds for every vector z in $\mathcal{A}$.

This result follows from Theorem 12.1 with $L(z) = P(x, z)$.

THEOREM 12.9. *Let $Q(x)$ be a quasi-nonsingular quadratic form and let $L_1(x)$,* $\cdots$, $L_p(x)$ *be p linearly independent linear forms on $\mathfrak{A}$. Denote by $\mathfrak{B}$ the set of vectors in $\mathfrak{A}$ satisfying the conditions*

$$(12.1) \qquad\qquad L_\alpha(x) = 0 \qquad\qquad (\alpha = 1, \cdots, p),$$

and denote by $\mathfrak{C}$ the Q-orthogonal complement of $\mathfrak{B}$. The set of vectors common to $\mathfrak{B}$ and $\mathfrak{C}$ is the class $\mathfrak{D}$ of all Q-transversals of $\mathfrak{A}$ satisfying (12.1). The class $\mathfrak{C}$ is the direct sum of $\mathfrak{D}$ and a linear class $\mathfrak{P}$ of dimension p orthogonal to $\mathfrak{D}$. If $\mathfrak{D}$ is of finite dimension d, then $\mathfrak{C}$ is of dimension $p + d$.

It is sufficient to consider the case where $\mathfrak{D}$ consists of the vector $x = 0$, since this can be brought about by replacing $\mathfrak{A}$ by the orthogonal complement of $\mathfrak{D}$. Then the nullity n of $Q(x)$ cannot exceed p. Let $x_1, \cdots, x_n$ be a basis for the class $\mathfrak{A}_0$ of Q-transversals of $\mathfrak{A}$. We can assume that the first $m = p - n$ linear forms $L_1(x), \cdots, L_m(x)$ vanish on $\mathfrak{A}_0$. Inasmuch as $Q(x)$ is quasi-nonsingular, there exist vectors $y_1, \cdots, y_m$ in $\mathfrak{A}$ such that

$$L_\beta(x) = Q(y_\beta, x) \qquad\qquad (\beta = 1, \cdots, m)$$

on $\mathfrak{A}$. The vectors $y_1, \cdots, y_m$ are in $\mathfrak{C}$, and the vectors $y_1, \cdots, y_m, x_1, \cdots, x_n$ form a linearly independent set in $\mathfrak{C}$. The dimension of $\mathfrak{C} = \mathfrak{P}$ is accordingly at least $p = m + n$. It cannot exceed p since $\mathfrak{D}$ is the zero class. This proves the theorem.

ILLUSTRATION 1. Consider Example II in §4 and let $J(x)$ be the quadratic form (4.13). Suppose the strengthened condition (4.9) of Legendre holds. Then, as was seen in the last section, $J(x)$ is a Legendre form and, by Theorem 12.6, is quasi-nonsingular on $\mathfrak{A}$. Let $\mathfrak{B}$ be the class of all arcs

$$x: \qquad\qquad x^j(t) \qquad\qquad (a \leq t \leq b, \ j = 1, \cdots, r)$$

in $\mathfrak{A}$ that vanish at $t = a$ and $t = b$. The J-orthogonal complement $\mathfrak{C}$ of $\mathfrak{B}$ is the class of extremals, that is, the solutions of equations (6.4). Its dimension is $2r$. The class $\mathfrak{B}_0$ of J-transversals of $\mathfrak{B}$ consists of all extremals that vanish at $t = a$ and $t = b$. The statement that $x = y + z$, where y is in $\mathfrak{B}$ and z is in $\mathfrak{C}$, is equivalent to the statement that the end-points of x can be joined by an extremal z. Thus, by virtue of Theorem 12.3, the end-points of an arc x in $\mathfrak{A}$ can be joined by an

extremal z if and only if x is J-orthogonal to the extremals that vanish at $t = a$ and $t = b$. To state this in another way, let $x_1, \cdots, x_s$ be a basis of $\mathcal{B}_0$. According to the result given in Illustration 2 in §6 there exist absolutely continuous functions $\xi_{j\alpha}(t)$ such that

$$(12.2) \qquad \dot{\xi}_{j\alpha}(t) = \omega_{\dot{x}^j}[t, x_\alpha(t), \dot{x}_\alpha(t)] \qquad (\alpha = 1, \cdots, s)$$

almost everywhere on $a \leq t \leq b$. Then, as is easily seen,

$$J(x_\alpha, x) = \xi_{j\alpha}(b)x^j(b) - \xi_{j\alpha}(a)x^j(a) .$$

Consequently, the end-points of an arc x in $\mathcal{A}$ can be joined by an extremal if and only if one has

$$\xi_{j\alpha}(b)x^j(b) = \xi_{j\alpha}(a)x^j(a) \qquad (\alpha = 1, \cdots, s).$$

It can be shown that the matrix $\|\xi_{j\alpha}(a)\|$ has rank s since, by virtue of the Legendre condition, no extremal $x \neq 0$ has $x^j(a) = \xi_j(a) = 0$, where ξ_j is the right member of (6.4). Consequently, if we interpret the conditions (12.1) to be the set $x^j(a) = 0$, $x^j(b) = 0$ ($j = 1, \cdots, r$), we see, by Theorem 12.9, that the dimension of the class $\mathcal{C}$ of extremals is $p = 2r$, a fact which is known from the theory of differential equations.

ILLUSTRATION 2. If in Illustration 1 we let $\mathcal{B}$ be the set of arcs in $\mathcal{A}$ which vanish at

$$t_0 = a < t_1 < \cdots < t_p < t_{p+1} = b ,$$

then the class $\mathcal{C}$ of vectors J-orthogonal to $\mathcal{B}$ consists of all broken extremals having corners at most at the points $t_1, \cdots, t_p$. Let $x_1, \cdots, x_s$ be a basis for the class $\mathcal{B}_0$ of all broken extremals in $\mathcal{B}$. Then there exist functions $\xi_{i\alpha}(t)$ that are absolutely continuous on each interval $t_{\beta-1} \leq t \leq t_\beta$ such that (12.2) holds almost everywhere on $a \leq t \leq b$. It is easily seen that

$$J(x_\alpha, x) = \xi_{k\alpha}(b)x^k(a) - \xi_{k\alpha}(a)x^k(a)$$

$$+ \sum_{\beta=1}^{p} [\xi_{k\alpha}(t_\beta - 0) - \xi_{k\alpha}(t_\beta + 0)]x^k(t_\beta) .$$

Given an arc x in $\mathcal{A}$ there exists a broken extremal y with corners at most at t_1, $\cdots$, t_p and having

$$y_k(t_\beta) = x_k(t_\beta) \qquad\qquad (\beta = 0, 1, \cdots, p + 1)$$

if and only if

$$J(x_\alpha, x) = 0 \qquad\qquad (\alpha = 1, \cdots, s).$$

By use of Theorem 12.9 and the remark made at the end of Illustration 1, it is seen that the dimension of the class $\mathcal{C}$ of broken extremals is $r(p + 1)$.

Further illustrations of this type can be found in the papers by Hazard and Ritcey, referred to in $\S 1$.

13. Legendre pairs. Two quadratic forms $P(x)$ and $Q(x)$ will be said to be a *Legendre pair* on $\mathcal{A}$ if (a) they are *wls*-continuous on $\mathcal{A}$; (b) $x_q \Longrightarrow x_0$ whenever $x_q \longrightarrow x_0$, $P(x_q) \longrightarrow P(x_0)$, $Q(x_q) \longrightarrow Q(x_0)$. It is easily seen that $P(x)$, $Q(x)$ form a Legendre pair if and only if

$$(13.1) \qquad\qquad J(x, b) = P(x) + bQ(x)$$

is a Legendre form for every positive number b. The corresponding bilinear form is

$$(13.2) \qquad\qquad J(x, y, b) = P(x, y) + bQ(x, y).$$

THEOREM 13.1. *If a Legendre pair $P(x)$, $Q(x)$ has the property that the relations $P(x) \le 0$, $Q(x) \le 0$ hold simultaneously only in case $x = 0$, then there is a positive number b such that the quadratic form (13.1) is positive definite on $\mathcal{A}$.*

The proof of the case where one of the forms is nonnegative will be left to the reader. Suppose therefore that neither $P(x)$ nor $Q(x)$ is nonnegative on $\mathcal{A}$. Let $\mathcal{S}$ be the set of points $x \ne 0$ for which $Q(x) \le 0$. Then by hypothesis $J(x, 0) = P(x) > 0$ on $\mathcal{S}$. Since $Q(x) < 0$ at some point in $\mathcal{S}$, there is a largest number b'' such that $J(x, b'') \ge 0$ on $\mathcal{S}$. We may select a sequence $\{y_q\}$ in $\mathcal{S}$ such that

$$|y_q| = 1, \quad J\left(y_q, b'' + \frac{1}{q}\right) \le 0, \quad Q(y_q) \le 0, \quad P(y_q) > 0,$$

the last two holding because y_q is in $\mathcal{S}$. In fact this sequence can be chosen to converge weakly to a vector y. By virtue of *wls*-continuity we have

$$|y| \le 1, \quad J(y, b'') \le 0, \quad Q(y) \le 0.$$

If $y = 0$ we have $y_q \longrightarrow 0$, $Q(y_q) \longrightarrow 0$, $P(y_q) \longrightarrow 0$, and hence $y_q \Longrightarrow 0$, contrary to the relation $|y_q| = 1$. Consequently y is a nonnull vector in $\mathcal{S}$. In view of our

choice of b'' we have $J(y, b'') = 0$. Since $P(y) > 0$ by virtue of our hypotheses, the relations

(13.3) $$b'' > 0, \quad Q(y) < 0, \quad P(y) > 0$$

must hold. It follows that given a vector x in $\hat{\mathbb{Q}}$ the vector $y + tx$ must be in $\hat{\mathscr{E}}$ for small values of t. Hence, for these values of t we have

$$0 \le J(y + tx, b'') = 2\, t\, J(y, x, b'') + t^2 J(x, b'') \,.$$

This is possible only in case the first two of the relations

(13.4) $$J(y, x, b'') = 0, \quad J(x, b'') \ge 0, \quad J(y, b'') = 0$$

hold. Thus b'' is the largest number such that $J(x, b'') \ge 0$ on $\hat{\mathbb{Q}}$.

Interchanging the roles of P and Q we see that there is a least positive number b' such that $J(x, b') \ge 0$ on $\hat{\mathbb{Q}}$. Moreover there is a vector z such that

(13.5) $$J(z, x, b') = 0, \quad J(z, b') = 0, \quad Q(z) > 0, \quad P(z) < 0$$

for all x in $\hat{\mathbb{Q}}$. Since $Q(y) < 0$ and $Q(z) > 0$, there is a vector $x \ne 0$ of the form $x = ay + cz$ such that $Q(x) = 0$. We have accordingly, by (13.4) and (13.5),

$$0 < P(x) = J(x, b'') = c^2 J(z, b'') = c^2 (b'' - b') Q(z) \,.$$

Hence $b'' > b'$. Moreover $J(x, b') = J(x, b'') = 0$ holds only in case $P(x) = Q(x) = 0$ and hence only in case $x = 0$. It follows that if b is on the interval $b' < b < b''$ and $x \ne 0$ we have, with $t = (b - b')/(b'' - b')$,

$$J(x, b) = (1 - t)\, J(x, b') + t\, J(x, b'') > 0 \,.$$

The form $J(x, b)$ is accordingly a positive Legendre form on $\hat{\mathbb{Q}}$ and hence is positive definite, as was to be proved.

The following lemma will be useful.

LEMMA 13.1. *If $P(x)$ and $Q(x)$ are quadratic forms such that $P(x) > 0$ whenever $x \ne 0$ and $Q(x) = 0$, then either $Q(x)$ or $-Q(x)$ is positive on the set of vectors $x \ne 0$ for which $P(x) \le 0$.*

For suppose there exist vectors x and y such that

$$P(x) \le 0, \quad Q(x) < 0, \quad P(y) \le 0, \quad Q(y) > 0 \,.$$

On the line $z = x + ty$ we have

$$P(z) = P(x) + 2t\, P(x, y) + t^2 P(y)\,,$$

$$Q(z) = Q(x) + 2t\, Q(x, y) + t^2 Q(y)\,.$$

Since $Q(x)\, Q(y) < 0$ there exist numbers $t' < 0 < t''$ such that $Q(z)$ vanishes at the corresponding points z' and z''. By virtue of our hypotheses we have $P(z') > 0$, $P(z'') > 0$. Since x lies between z' and z'', the equation $P(z) = 0$ in t has two roots on $t't''$. This is impossible if $P(y) = 0$. Hence $P(y) < 0$ and $P(z)$ would vanish also exterior to $t't''$. This too is impossible. In view of this contradiction the lemma is established.

THEOREM 13.2. *Let $J(x)$ be a Legendre form and let $K(x)$ be a w-continuous form with the property that $J(x) > 0$ whenever $K(x) = 0$ and $x \neq 0$. Then there is a number b such that $J(x) + bK(x)$ is positive definite on $\mathfrak{A}$. If $K(x) \geq 0$ whenever $J(x) \leq 0$, the number b can be chosen to be positive.*

In view of Lemma 12.1, either $K(x)$ or $-K(x)$ is positive whenever $J(x) \leq 0$ and $x \neq 0$. We can suppose $K(x)$ has this property. Then $P(x) = J(x)$, $Q(x) = K(x)$ have the properties prescribed in Theorem 13.1, and the theorem follows.

The result described in the last two theorems is a generalization of results given by several authors: see [1], [9], [16], and [20].

THEOREM 13.3. *Let $Q(x)$ be a nonnegative quasi-nonsingular quadratic form on $\mathfrak{A}$, and let $P(x)$ be a Legendre form on the class $\mathfrak{B}$ of Q-transversals of $\mathfrak{A}$. Then $P(x)$ and $Q(x)$ form a Legendre pair on $\mathfrak{A}$. Moreover there is a positive number b such that*

(13.6) $$J(x) = P(x) + bQ(x)$$

is a Legendre form on $\mathfrak{A}$ having the same index and nullity on $\mathfrak{A}$ as those of $P(x)$ on $\mathfrak{B}$.

To show that $P(x)$ and $Q(x)$ form a Legendre pair, suppose $x_q \longrightarrow x_0$, $P(x_q) \longrightarrow P(x_0)$, $Q(x_q) \longrightarrow Q(x_0)$. Select vectors y_q, y_0 in $\mathfrak{B}$ and z_q, z_0 in the orthogonal complement $\mathfrak{C}$ of $\mathfrak{B}$ such that $x_q = y_q + z_q$, $x_0 = y_0 + z_0$. Then $y_q \longrightarrow y_0$, $z_q \longrightarrow z_0$. Moreover $Q(x_q) = Q(z_q) \longrightarrow Q(x_0) = Q(z_0)$. Since $Q(x)$ is positive definite on $\mathfrak{C}$, by Theorem 12.4, it follows that $z_q \Longrightarrow z_0$. We have accordingly $P(y_q, z_q) \longrightarrow P(y_0, z_0)$, $P(z_q) \longrightarrow P(z_0)$. But this implies that $P(y_q) \longrightarrow P(y_0)$ and hence that $y_q \Longrightarrow y_0$, since $P(x)$ is a Legendre form on $\mathfrak{B}$. Consequently, $x_q \Longrightarrow x_0$, as was to be proved.

In order to prove the last statement in the theorem, let i be the index and n the nullity of $P(x)$ on $\mathfrak{B}$. Choose a linear subclass $\mathfrak{D}$ of $\mathfrak{B}$ of dimension $i + n$ such that $P(x)$ is positive for all vectors $x \neq 0$ in $\mathfrak{B}$ orthogonal to $\mathfrak{D}$. Then the hypotheses of Theorem 13.1 hold on the orthogonal complement $\mathfrak{C}^*$ of $\mathfrak{D}$. There is accordingly a positive constant b such that the quadratic form $J(x)$ defined by (13.6) is positive definite on $\mathfrak{C}^*$. Consequently, $J(x)$ is a Legendre form on $\mathfrak{C}$, the sum of whose index and nullity is at most $i + n$. Since $J(x) = P(x)$ on $\mathfrak{B}$, its index and nullity must be given by i and n respectively.

14. The Lagrange multiplier rule. It is of interest to pause for a moment to show that the Lagrange multiplier rule can be obtained by an application of the concept of quasi-nonsingularity. To this end we consider the class $\mathfrak{C}$ of arcs described in §4. Let $\mathfrak{B}$ be the class of all arcs x in $\mathfrak{C}$ satisfying a set of differential equations

$$(14.1) \quad \phi_\sigma(t, x, \dot{x}) = M_{\sigma k}(t)x^k(t) + N_{\sigma k}(t)\dot{x}^k(t) = 0 \qquad (\sigma = 1, \cdots, r < p)$$

almost everywhere on $a \leq t \leq b$. We shall assume that the coefficients $M_{\sigma k}(t)$ and $N_{\sigma k}(t)$ are continuous on $a \leq t \leq b$ and that the matrix $\|N_{\sigma k}(t)\|$ $(\sigma = 1, \cdots, r; k = 1, \cdots, p)$ has rank r on this interval. As is well known, there exist functions

$$(14.2) \quad \phi_\tau(t, x, \dot{x}) = M_{\tau k}(t)x^k + N_{\tau k}(t)\dot{x}^k \qquad (\tau = r + 1, \cdots, p)$$

with continuous coefficients such that the matrix $\|N_{jk}(t)\|$ $(j, k = 1, \cdots, p)$ has rank p on $a \leq t \leq b$. Then the relations

$$N_{ij}(t) N_{ik}(t) \pi^j \pi^k > 0 \qquad (i, j, k = 1, \cdots, p)$$

hold for every set $(\pi) \neq (0)$. The quadratic form $D(x) = D(x, x)$, where

$$D(x, y) = x^k(a)y^k(a) + \int_a^b \phi_k[t, x(t), \dot{x}(t)] \, \phi_k[t, y(t), \dot{y}(t)] \, dt \,,$$

is accordingly (see §11) a nonnegative Legendre form. Moreover $D(x) = 0$ if and only if $x^k(a) = 0$ and $\phi_k[t, x(t), \dot{x}(t)] = 0$, that is, if and only if $x = 0$. The form $D(x)$ is accordingly a positive definite form on $\mathfrak{C}$. With the help of $D(x)$ we can prove the following:

THEOREM 14.1. *Let $Q(x)$ be the quadratic form*

$$(14.3) \quad Q(x) = L_\alpha(x)L_\alpha(x) + \int_a^b \phi_\sigma(t, x, \dot{x}) \, \phi_\sigma(t, x, \dot{x}) \, dt \,,$$

where $\alpha = 1, \cdots, m$; $\sigma = 1, \cdots, r$; *and* $L_1(x), \cdots, L_m(x)$ *are linear forms. Then the quadratic form* $Q(x)$ *is a nonnegative quasi-nonsingular form on* $\mathbb{G}$.

In view of the corollary to Theorem 12.5, it is sufficient to consider the case $m = 0$. Then the class of arcs on which $Q(x) = 0$ is the class $\mathbb{B}$ of arcs on which $\phi_\sigma = 0$. It follows that if x is in $\mathbb{B}$ then

$$D(x, y) = x^k(a)y^k(a) + \int_a^b \phi_\tau(t, x, \dot{x}) \, \phi_\tau(t, y, \dot{y}) \, dt \qquad (\tau = r + 1, \cdots, p).$$

Since an arc x in $\mathbb{B}$ is completely determined when the values $x^k(a)$ and $\phi_\tau(t) = \phi_\tau[t, x(t), \dot{x}(t)]$ are given, it follows that an arc y is D-orthogonal to $\mathbb{B}$ if and only if $y^k(a) = 0$ and $\phi_\tau[t, y(t), \dot{y}(t)] = 0$ $(\tau = r + 1, \cdots, p)$. On the class $\mathbb{C}$ of D-orthogonals of $\mathbb{B}$ we have accordingly $D(y) = Q(y)$. The form $Q(y)$ is accordingly positive definite on $\mathbb{C}$. In view of Theorem 12.4, with $D(x, y)$ as the inner product, the quadratic form $Q(x)$ is quasi-nonsingular on $\mathbb{G}$, as was to be proved.

Consider now the case

$$(14.4) \qquad L_\alpha(x) = a_{\alpha k} x^k(a) + b_{\alpha k} x^k(b) + \int_a^b \omega_\alpha(t, x, \dot{x}) \, dt \,,$$

where ω_α is linear in x^k and $\dot{x}^k$, the coefficients being continuous functions of t. We are now in position to prove the following:

THEOREM 14.2. LAGRANGE MULTIPLIER RULE. *Let* $\mathbb{G}^*$ *be the class of arcs in* $\mathbb{G}$ *satisfying the conditions*

$$(14.5) \qquad L_\alpha(x) = 0, \quad \phi_\sigma(t, x, \dot{x}) = 0 \qquad (\alpha = 1, \cdots, m; \, \sigma = 1, \cdots, r),$$

where L_α *and* ϕ_σ *are given by* (14.4) *and* (14.1), *respectively. Let* $J(x)$ *be the quadratic form* (4.13). *Then a vector* x *in* $\mathbb{G}$ *is* J-*orthogonal to* $\mathbb{G}^*$ *if and only if there exist constants* c_k *and a function*

$$\Omega = \omega + \mu_\alpha \omega_\alpha + \lambda_\sigma(t) \, \phi_\sigma \,,$$

where μ_α *are constants and* $\lambda_\sigma(t)$ *are square integrable functions, such that* (a) *the Euler-Lagrange equations*

$$(14.6) \qquad \Omega_{\dot{x}^k} = \int_a^t \Omega_{x^k} \, dt + c_k$$

hold almost everywhere on $a \leq t \leq b$; (b) *the functions* $\xi_k(t)$ *defined by the right*

member of (14.6) *satisfy the transversality conditions*

(14.7) $$-\xi_k(a) + \mu_\alpha a_{\alpha k} + q_{ka}(x) = 0,$$

$$\xi_k(b) + \mu_\alpha b_{\alpha k} + q_{kb}(x) = 0,$$

where q_{ka}, q_{kb} *have the significance described in Illustration 2 of* §6.

In order to prove this fact recall that the quadratic form $Q(x)$ defined by (14.3) is quasi-nonsingular on $\mathcal{Q}$. Moreover, $Q(x) = 0$ if and only if x is in $\mathcal{Q}^*$. Thus, $\mathcal{Q}^*$ is the class of Q-transversals of $\mathcal{Q}$. If a vector x is J-orthogonal to $\mathcal{Q}^*$, then the linear form $L(z) = J(x, z)$ is zero on the set of Q-transversals of $\mathcal{Q}$. There is accordingly a vector y in $\mathcal{Q}$ such that

(14.8) $$J(x, z) + Q(y, z) = 0$$

for all z in $\mathcal{Q}$. If we set

$$\mu_\alpha = L_\alpha(y), \quad \lambda_\sigma(t) = \phi_\sigma[t, y(t), \dot{y}(t)],$$

equation (14.8) takes the form

$$[q_{ka}(x) + \mu_\alpha a_{\alpha k}(x)]z^k(a) + [q_{ka}(x) + \mu_\alpha b_{\alpha k}(x)]z^k(b)$$

$$+ \int_a^b \left\{ \Omega_{x^k} z^k + \Omega_{\dot{x}^k} \dot{z}^k \right\} dt = 0,$$

where Ω is defined as described in the theorem. Using Theorem 4.5, we obtain (14.6) and (14.7). Retracing our steps, we see that the converse also is true.

In view of the following theorem the criterion that $J(x)$ be a Legendre form on $\mathcal{Q}^*$ is equivalent to the strengthened condition of Legendre-Clebsch for problems of Lagrange and Bolza.

THEOREM 14.3. *Let* $J(x)$ *and* $\mathcal{Q}^*$ *be defined as in Theorem 14.2. Then the quadratic form* $J(x)$ *is a Legendre form on* $\mathcal{Q}^*$ *if and only if there is a number* $h > 0$ *such that at almost all points* t *on* $a \leq t \leq b$ *the inequality*

(14.9) $$R_{jk}(t) \pi^j \pi^k \geq h \pi^k \pi^k$$

holds for every solution $(\pi) \neq (0)$ *of*

(14.10) $$N_{\alpha k}(t)\pi^k = 0 \qquad\qquad (\alpha = 1, \cdots, r).$$

For by the use of Theorem 12.3 with $P(x) = J(x)$, and $Q(x)$ given by (14.3), the

quadratic form $J(x)$ is a Legendre form on $\mathcal{C}^*$ if and only if there is a constant $\lambda > 0$ such that $I(x) = J(x) + \lambda Q(x)$ is a Legendre form on $\mathcal{C}$. In view of Illustration 1 of §11, the form $I(x)$ is a Legendre form on $\mathcal{C}$ if and only if at almost all points of $a \leq t \leq b$ the inequality

$$R_{jk}\pi^j\pi^k + \lambda N_{\alpha j}N_{\alpha k}\pi^j\pi^k \geq h\pi^k\pi^k$$

holds for every set $\pi \neq (0)$ for a suitably chosen positive number h. By virtue of Theorem 13.3 for the finite dimensional case, this condition is equivalent to condition (14.9) subject to the restriction (14.10), as was to be proved.

The above remarks together with Theorem 13.3 yield the following consequent result.

COROLLARY. *If $J(x)$ is a Legendre form on $\mathcal{C}^*$ of index i and nullity n, then there is a Legendre form $I(x)$ on $\mathcal{C}$ of index i and nullity n that coincides with $J(x)$ on $\mathcal{C}^*$.*

15. Quasi-Legendre forms. As was previously seen in §11, the quadratic forms that arise in the study of nonparametric problems satisfy the strengthened condition of Legendre if and only if they are Legendre forms. This is not true for the quadratic forms $J(x)$ that arise in the study of parametric problems. The difficulty lies in the fact that in this case the class $\mathcal{C}_0$ of J-transversals of $\mathcal{C}$ is of infinite dimension, which is impossible for Legendre forms. It can be shown (but we shall not do so here) that in this case $J(x)$ will be a Legendre form on the orthogonal complement of $\mathcal{C}_0$ if the strengthened condition of Legendre holds. Such forms belong to a class that we shall call quasi-Legendre forms. The definition we shall adopt is the following one:

A quadratic form $Q(x)$ will be called a *quasi-Legendre form* on $\mathcal{C}$ in case it is *wls*-continuous and quasi-nonsingular.

In view of Theorem 12.7 we have:

THEOREM 15.1. *A quadratic form $Q(x)$ is a quasi-Legendre form on $\mathcal{C}$ if and only if it is a Legendre form on the orthogonal complement of the Q-transversals of $\mathcal{C}$. Moreover, $Q(x)$ is a quasi-Legendre form if and only if it is quasi-nonsingular and of finite index on $\mathcal{C}$.*

As an application of quasi-nonsingularity to the theory of indices we have the following:

THEOREM 15.2. *Let $Q(x)$ be a quasi-Legendre form on $\mathcal{C}$ and denote by $\mathcal{C}_0$ the*

set of Q-transversals of $\mathfrak{A}$. Let $\mathfrak{A}_1$ be a closed linear subspace of $\mathfrak{A}$ and denote by $\mathfrak{A}_2$ the Q-orthogonal complement of $\mathfrak{A}_1$. Finally, let C be the maximal linear class of vectors common to $\mathfrak{A}_1$ and $\mathfrak{A}_2$, having no nonnull vector in common with $\mathfrak{A}_0$. Then the index i of $Q(x)$ on $\mathfrak{A}$ is given by the formula

$$(15.1) \qquad\qquad i = i_1 + i_2 + c,$$

where i_1, i_2 are the indices of $Q(x)$ on $\mathfrak{A}_1$, $\mathfrak{A}_2$ respectively, and c is the dimension of C. If $Q(x)$ is of finite nullity n on $\mathfrak{A}$, and n_1, n_2 are the nullities of $Q(x)$ on $\mathfrak{A}_1$ and $\mathfrak{A}_2$, then

$$(15.2) \qquad\qquad n_2 = n + c = n_1 + p,$$

where p is the dimension of a maximal class $\mathfrak{P}$ of Q-transversals of $\mathfrak{A}$ having no $x \neq 0$ in common with $\mathfrak{A}_1$.

In order to prove this result, let $\mathfrak{B}_1$ and $\mathfrak{B}_2$ be maximal linear subclasses of $\mathfrak{A}_1$ and $\mathfrak{A}_2$ on which $Q(x)$ is negative. Their dimensions are i_1 and i_2, respectively. Let $\mathfrak{B}$ be the direct sum of the classes $\mathfrak{B}_1$, $\mathfrak{B}_2$, and C. Its dimension is $k = i_1 + i_2 + c$. Clearly $Q(x) \leq 0$ on $\mathfrak{B}$. Moreover, no vector $x \neq 0$ in $\mathfrak{B}$ is in $\mathfrak{A}_0$. Consider a vector x that is Q-orthogonal to $\mathfrak{B}$. The vector x is Q-orthogonal to C and hence to the class of Q-transversals of $\mathfrak{A}_1$. By Theorem 12.3, there are vectors x_1 in $\mathfrak{A}_1$ and x_2 in $\mathfrak{A}_2$ such that $x = x_1 + x_2$. Inasmuch as x and x_1 are Q-orthogonal to $\mathfrak{B}_2$, so also is x_2. Hence $Q(x_2) \geq 0$, the equality holding only in case x_2 is a Q-transversal of $\mathfrak{A}_2$, by Lemma 9.1. Similarly, we have $Q(x_1) \geq 0$, the equality holding only in case x_1 is a Q-transversal of $\mathfrak{A}_1$ and hence also of $\mathfrak{A}_2$. It follows that $Q(x) = Q(x_1) + Q(x_2) \geq 0$, the equality holding if and only if x is a Q-transversal of $\mathfrak{A}_2$. From this result we see that $\mathfrak{B}$ is a maximal class on which $Q(x) \leq 0$ and which does not contain a Q-transversal of $\mathfrak{A}$. The dimension of $\mathfrak{B}$ is accordingly equal to the index i of $Q(x)$ on $\mathfrak{A}$, by Theorem 9.1. The last statement in the theorem follows from the definition of C.

The result obtained in Theorem 15.2 can be restated in the manner described in the following:

THEOREM 15.3. *Let $Q(x)$ be a quasi-Legendre form on $\mathfrak{A}$ and denote by $\mathfrak{A}_1$ a closed linear subspace of $\mathfrak{A}$. Let i, i_1 be the indices of $Q(x)$ on $\mathfrak{A}$, $\mathfrak{A}_1$, respectively. The difference $i - i_1$ is equal to the dimension of a maximal linear subclass $\mathfrak{D}$ of $\mathfrak{A}$ such that (a) $\mathfrak{D}$ is Q-orthogonal to $\mathfrak{A}_1$; (b) $Q(x) \leq 0$ on $\mathfrak{D}$; (c) $\mathfrak{D}$ contains no nonnull Q-transversal of $\mathfrak{A}$. If n, n_1 are the nullities of $Q(x)$ on $\mathfrak{A}$,*

$\mathcal{C}_1$, and $m = i + n$, $m_1 = i_1 + n_1$, then $m = m_1 + e$, where e is the dimension of a maximal linear subclass $\mathcal{E}$ of $\mathcal{C}$ such that (a) the subclass $\mathcal{E}$ is Q-orthogonal to $\mathcal{C}_1$; (b) the subclass $\mathcal{E}$ has no element $x \neq 0$ in common with $\mathcal{C}_1$; (c) we have $Q(x) \leq 0$ on $\mathcal{E}$.

This result follows from the proof of Theorem 15.2.

The theorems just established have many applications in the calculus of variations. We shall describe only a few of these. These results also hold when the class $\mathcal{C}$ is replaced by the class $\mathcal{C}^*$ described in Theorem 14.2. Similar results hold for multiple integral problems.

In the following illustrations, $\mathcal{C}$ is the class described in §4, and $J(x)$ is a Legendre form on $\mathcal{C}$.

ILLUSTRATION. 1. Let $\mathcal{C}_1$ be the subclass of all arcs in $\mathcal{C}$ having $x(a) = x(b) = 0$. The class $\mathcal{E}$ of arcs J-orthogonal to $\mathcal{C}_1$ is the class of extremal arcs. Let $\mathcal{C}^*$ be a linear subclass of $\mathcal{C}$ containing $\mathcal{C}_1$. Applying Theorem 15.2 with $\mathcal{C}^*$ playing the role of $\mathcal{C}$, we see that the index i^* of $J(x)$ on $\mathcal{C}^*$ is given by the sum

$$\iota^* = i_1 + i_2 + c ,$$

where i_1 is the index $\mathcal{C}_1$, i_2 is the dimension of a maximal class of extremals in $\mathcal{C}^*$ on which $J(x)$ is negative, and c is the dimension of the class of extremals belonging to $\mathcal{C}_1$ and $\mathcal{C}_2$.

Before proceeding to the next illustration we find it convenient to establish a lemma.

LEMMA 15.1. *There is a constant $\delta > 0$ such that $J(x) > 0$ for every arc $x \neq 0$ in $\mathcal{C}$ vanishing somewhere on every subinterval of $a \leq t \leq b$ of length δ.*

For if this were not the case there would exist a set of vectors x_q, x_0 satisfying the conditions

$$J(x_q) \leq 0, \quad |x_q| = 1, \quad x_q \longrightarrow x_0 ,$$

such that x_q vanishes somewhere on every subinterval of $a \leq t \leq b$ of length $1/q$. The vector x_0 is obviously the zero vector. Since $J(x)$ is *wls*-continuous, we have $J(x_q) \longrightarrow J(x_0) = 0$, and hence also the relations $x_q \Longrightarrow x_0 = 0$, contrary to the choice of the x_q as unit vectors.

ILLUSTRATION 2. Let $t_1 < \cdots < t_r$ be points on $a < t < b$ which divide

$a \leq t \leq b$ into subintervals of length $\leq \delta$, where δ has the property described in Lemma 15.1. Let $\mathcal{G}_1$ be the class of arcs in $\mathcal{G}$ which vanish at $t_0 = a, t_1, \cdots,$ $t_r, t_{r+1} = b$. Then $J(x)$ is positive definite on $\mathcal{G}$, by Lemma 16.1. The class of arcs $\mathcal{E}$ that are J-orthogonal to $\mathcal{G}_1$ is the class of broken extremals with corners at $t_1, \cdots, t_r$. Let $\mathcal{G}^*$ be a closed linear subclass of $\mathcal{G}$ containing $\mathcal{G}_1$, and let $\mathcal{G}_2$ be the class of broken extremals in $\mathcal{G}^*$. Applying Theorem 15.2 with $\mathcal{G}$ replaced by $\mathcal{G}^*$, we see that the index i^* of $J(x)$ on $\mathcal{G}^*$ is equal to the index i_2 of $J(x)$ on the class $\mathcal{G}_2$ of broken extremals in $\mathcal{G}^*$. This illustration identifies the concept of index given here with that given by Morse [13].

ILLUSTRATION 3. Let $\mathcal{G}^*$ be the class of all arcs in $\mathcal{G}$ that vanish at $t = b$, and denote by $\mathcal{G}_1$ the class of arcs in $\mathcal{G}^*$ that vanish on an interval $t_0 \leq t \leq b$, where $t_0 > a$. Let $\mathcal{G}_2$ be the class of arcs in $\mathcal{G}^*$ that are J-orthogonal to $\mathcal{G}_1$. If $J(x)$ is of the form (4.13), then an arc x in $\mathcal{G}^*$ is in $\mathcal{G}_2$ if and only if the segment determined by $a \leq t \leq t_0$ is a solution of the Euler equations satisfying the transversality condition at $t = a$. It follows that in this case an arc $x \neq 0$ belonging to $\mathcal{G}_1$ and $\mathcal{G}_2$ exists if and only if there is an extremal satisfying the transversality condition at $t = a$ and vanishing at $t = t_0$. The number of linearly independent extremals of this type, no linear combination of which is identically zero on $t_0 \leq t \leq b$, is equal to the number c described in Theorem 16.2. As we shall presently see, c is the order of t_0 as a focal point of our problem. By virtue of Lemma 16.1 we have $c = 0$ and $i_1 = 0$ when $t_0 \leq a + \delta$.

16. Focal points. The theory of focal points and conjugate points in the calculus of variations can be extended so as to be applicable to a quasi-Legendre form $J(x)$ on a Hilbert space $\mathcal{G}$. Connections between the results here developed and those found in the calculus of variations will be given in the next section.

In the present section it will be assumed that we have given a *one-parameter family of closed linear subclasses* $\mathcal{G}(\lambda)$ ($\lambda' \leq \lambda \leq \lambda''$) *of* $\mathcal{G}$ *such that* (a) $\mathcal{G}(\lambda')$ *is the zero class and* $\mathcal{G}(\lambda'') = \mathcal{G}$; (b) $\mathcal{G}(\lambda_1)$ *is a subclass of* $\mathcal{G}(\lambda_2)$ *whenever* $\lambda_1 < \lambda_2$; (c) *if* $\lambda' \leq \lambda_0 < \lambda''$ *then the class* $\mathcal{G}(\lambda_0)$ *is the intersection of all classes* $\mathcal{G}(\lambda)$ *with* $\lambda_0 < \lambda \leq \lambda''$.

Normally the classes in which we shall be interested have the additional property that (d) if $\lambda' < \lambda_0 \leq \lambda''$, then $\mathcal{G}(\lambda_0)$ is identical with the closure $\mathcal{G}(\lambda_0 - 0)$ of the union of the classes $\mathcal{G}(\lambda)$ with $\lambda' \leq \lambda < \lambda_0$. However, this property does not hold in a number of important applications and so we shall not make this further restriction here.

One should observe in passing that the classes $\mathcal{A}(\lambda)$ defined above correspond to a resolution of the identity. In other words, corresponding to each resolution of the identity there is a theory of focal points.

LEMMA 16.1. *If* $\lambda' < \lambda \leq \lambda''$, *then the left-hand limit* $i(\lambda - 0)$ *is equal to the index of* $J(x)$ *on the class* $\mathcal{A}(\lambda - 0)$ *described above. In particular if* $\mathcal{A}(\lambda - 0)$ $= \mathcal{A}(\lambda)$, *then* $i(\lambda - 0) = i(\lambda)$.

Consider a value λ_0 on $\lambda' < \lambda \leq \lambda''$, and let h be the index of $J(x)$ on $\mathcal{A}(\lambda_0 - 0)$. Clearly $i(\lambda) \leq h$ if $\lambda < \lambda_0$. We shall show that there is a value $\lambda_1 < \lambda_0$ such that $i(\lambda_1) = h_0$. This is obvious if $h = 0$. Suppose therefore that $h > 0$, and let x_1, $\cdots, x_h$ be a basis of a maximal linear subclass $\mathcal{B}$ of $\mathcal{A}(\lambda_0 - 0)$ on which $J(x) < 0$. For each integer $p \leq h$ select a sequence $\{x_{pr}\}$ in the union of the classes $\mathcal{A}(\lambda)$ $(\lambda' \leq \lambda < \lambda_0)$ converging strongly to x_p. We have accordingly

$$\lim_{r=\infty} J(x_{pr}, x_{qr}) a_p a_q = J(x_p, x_q) a_p a_q < 0 \qquad (p, \; q = 1, \; \cdots, \; h)$$

for every set $a_1, \cdots, a_h$ not all zero. Consequently there is an integer r such that

$$J(x_{pr}, x_{qr}) a_p a_q < 0$$

unless the a's are all zero. Choose $\lambda_1 < \lambda_0$ such that $x_{1r}, \cdots, x_{hr}$ are in $\mathcal{A}(\lambda_1)$. In view of Theorem 9.1 we have $i(\lambda_1) \geq h$ and hence $i(\lambda_1) = h$, as was to be proved.

LEMMA 16.2. *Let* $\mathcal{B}(\lambda)$ $(\lambda' < \lambda \leq \lambda'')$ *be a maximal linear subspace of* $\mathcal{A}(\lambda)$ *such that* (a) $\mathcal{B}(\lambda)$ *is* J-*orthogonal to* $\mathcal{A}(\lambda_1)$ *if* $\lambda_1 < \lambda$; (b) $J(x) \leq 0$ *on* $\mathcal{B}(\lambda)$; (c) *no* $x \neq 0$ *in* $\mathcal{B}(\lambda)$ *is* J-*orthogonal to* $\mathcal{A}(\lambda)$. *Then we have*

$$i(\lambda) = i(\lambda - 0) + b(\lambda),$$

where $b(\lambda)$ *is the dimension of* $\mathcal{B}(\lambda)$. *If* $\mathcal{A}(\lambda - 0) = \mathcal{A}(\lambda)$ $(\lambda' < \lambda \leq \lambda'')$, *then* $b(\lambda) = 0$ $(\lambda' < \lambda \leq \lambda'')$.

This result follows Lemma 16.1 and Theorem 15.3 since $\mathcal{B}(\lambda)$ is J-orthogonal to the class $\mathcal{A}(\lambda - 0)$.

LEMMA 16.3. *On the interval* $\lambda' \leq \lambda < \lambda''$, *the right-hand limit* $i(\lambda + 0)$ *is given by the formula*

$$i(\lambda + 0) = i(\lambda) + c(\lambda),$$

where $c(\lambda)$ is the dimension of a maximal linear subclass $\mathbb{C}(\lambda)$ of the J-transversals of $\mathbb{G}(\lambda)$ with the property that no $x \neq 0$ in $\mathbb{C}(\lambda)$ is J-orthogonal to a class $\mathbb{G}(\lambda_1)$ with $\lambda_1 > \lambda$. In particular, we have $i(\lambda' + 0) = 0$.

It is easily seen that two classes having the properties of $\mathbb{C}(\lambda)$ are of the same dimension. Moreover, by Theorem 9.1, we have $i(\lambda) + c(\lambda) \leq i(\lambda_1)$ if $\lambda < \lambda_1$. There is accordingly at most a finite number of points at which $c(\lambda) \neq 0$. Hence, given value λ_0 on $\lambda' \leq \lambda < \lambda''$, we may select λ_1 so that $c(\lambda) = 0$ on $\lambda_0 < \lambda \leq \lambda_1$. Let $\mathbb{D}(\lambda_0)$ be the class of J-transversals of λ_0 that are also J-transversals of some $\mathbb{G}(\lambda)$ with $\lambda > \lambda_0$. By virtue of our choice of λ_1, the vectors in $\mathbb{D}(\lambda_0)$ are J-transversals of $\mathbb{G}(\lambda)$ $(\lambda_0 \leq \lambda \leq \lambda_1)$. We can suppose that $\mathbb{C}(\lambda_0)$ is orthogonal to $\mathbb{D}(\lambda_0)$. Let $\mathbb{B}$ be a maximal linear subclass of $\mathbb{G}(\lambda_0)$ on which $J(x)$ is negative. Its dimension is $i(\lambda_0)$. Let $\mathbb{P}(\lambda)$ $(\lambda_0 \leq \lambda \leq \lambda_1)$ be the class of vectors in $\mathbb{G}(\lambda)$ that are J-orthogonal to $\mathbb{B}$ and orthogonal to $\mathbb{C}(\lambda_0)$ and $\mathbb{D}(\lambda_0)$. There is a value λ_2 on $\lambda_0 < \lambda \leq \lambda_1$ such that $J(x)$ is positive on $\mathbb{P}(\lambda)$. If this were not so one could find for every integer r a vector in $\mathbb{P}(\lambda_0 + 1/r)$ such that $|x_r| = 1$, $J(x_r) \leq 0$. The sequence $\{x_r\}$ can be chosen so as to converge weakly to the vector x_0. It is clear from property (c) of the family $\mathbb{G}(\lambda)$ that x_0 is in $\mathbb{G}(\lambda_0)$ and hence in $\mathbb{P}(\lambda_0)$. Consequently $J(x_0) \geq 0$, the equality holding only in case $x_0 = 0$. Inasmuch as $J(x)$ is *wls*-continuous, we have

$$0 \geq \lim_{r=\infty} \sup J(x_r) \geq \lim_{r=\infty} \inf J(x_r) \geq J(x_0) \geq 0.$$

It follows that $J(x_0) = 0$, and hence that $x_0 = 0$. We have accordingly $x_r \longrightarrow 0$, $J(x_r) \longrightarrow 0$. Since $J(x)$ is a Legendre form on $\mathbb{P}(\lambda_1)$, this result holds only in case $x_r \Longrightarrow 0$, contrary to the relation $|x_r| = 1$. This proves Lemma 16.3.

COROLLARY. *If there is no vector $x \neq 0$, which is simultaneously a J-transversal of $\mathbb{G}(\lambda_1)$ and of $\mathbb{G}(\lambda_2)$ for distinct values λ_1 and λ_2 on $\lambda' \leq \lambda \leq \lambda''$, then*

$$i(\lambda + 0) = i(\lambda) + n(\lambda) \qquad\qquad (\lambda' \leq \lambda < \lambda''),$$

where $n(\lambda)$ is the nullity of $J(x)$ on $\mathbb{G}(\lambda)$.

A point λ at which $i(\lambda)$ is discontinuous will be called a *focal point* of $J(x)$ relative to $\mathbb{G}(\lambda)$ $(\lambda' \leq \lambda \leq \lambda'')$, and the difference $f(\lambda) = i(\lambda + 0) - i(\lambda - 0)$ will be called the *order* of λ as a focal point. We set $i(\lambda'' + 0) = i(\lambda'')$ and $i(\lambda' - 0) = i(\lambda')$. Clearly $f(\lambda') = 0$, by Lemma 16.3.

Combining the results described in the last three lemmas we obtain:

THEOREM 16.1. *The order $f(\lambda)$ of λ as a focal point of $J(x)$ relative to $\mathcal{C}(\lambda)$ $(\lambda' \leq \lambda \leq \lambda'')$ is given by the formula*

$$f(\lambda) = b(\lambda) + c(\lambda),$$

where $b(\lambda)$ is defined in Lemma 16.2 with $b(\lambda') = 0$, and $c(\lambda)$ is defined in Lemma 16.3 with $c(\lambda'') = 0$. We have $f(\lambda') = 0$. If $\lambda' < \lambda \leq \lambda''$, then $f(\lambda)$ is the dimension of a maximal linear subclass $\mathfrak{J}(\lambda)$ of $\mathcal{C}(\lambda)$ such that (a) $\mathfrak{J}(\lambda)$ is J-orthogonal to $\mathcal{C}(\lambda_1)$ if $\lambda_1 < \lambda$, (b) $J(x) \leq 0$ on $\mathfrak{J}(\lambda)$, (c) no $x \neq 0$ in $\mathfrak{J}(\lambda)$ is J-orthogonal to a class $\mathcal{C}(\lambda_2)$ with $\lambda_2 > \lambda$ if $\lambda < \lambda''$ and with $\lambda_2 = \lambda''$ if $\lambda = \lambda''$.

The last statement can be proved directly with the help of Lemma 16.3.

COROLLARY. *If $\mathcal{C}(\lambda - 0) = \mathcal{C}(\lambda)$ $(\lambda' < \lambda \leq \lambda'')$ and if no vector $x \neq 0$ is J-orthogonal to $\mathcal{C}(\lambda_1)$ and $\mathcal{C}(\lambda_2)$ when $\lambda_1 \neq \lambda_2$, then $f(\lambda') = f(\lambda'') = 0$, and $f(\lambda)$ is the nullity of $J(x)$ on $\mathcal{C}(\lambda)$ $(\lambda' < \lambda < \lambda'')$.*

Oscillation theorems for differential equations of the second order can be established with the help of the following:

THEOREM 16.2. *Let $\mathcal{C}^*$ be a closed linear subspace of $\mathcal{C}$. Denote by $\mathcal{C}^*(\lambda)$ $(\lambda' \leq \lambda \leq \lambda'')$ the set of all vectors in $\mathcal{C}^*$ belonging to $\mathcal{C}(\lambda)$. The classes $\mathcal{C}^*(\lambda)$ have the properties (a), (b), (c) described at the beginning of this section with $\mathcal{C}$ replaced by $\mathcal{C}^*$. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ be the focal points of $J(x)$ relative to $\mathcal{C}(\lambda)$ $(\lambda' \leq \lambda \leq \lambda'')$, and let $\lambda_1^* \leq \lambda_2^* \leq \cdots \leq \lambda_{m^*}^*$ be the focal points of $J(x)$ relative to $\mathcal{C}^*(\lambda)$ $(\lambda' \leq \lambda \leq \lambda'')$, each repeated a number of times equal to its order. Then the relations*

(16.1) $$\lambda_r \leq \lambda_r^* \qquad\qquad (r = 1, \cdots, m^*)$$

hold. If $\mathcal{C}^*$ is the orthogonal complement of a linear subclass $\mathfrak{D}$ of dimension k, then in addition

(16.2) $$\lambda_r^* \leq \lambda_{r+k}$$

holds provided $r + k \leq m$.

The inequality (16.1) follows from the fact that the indices $i(\lambda)$ and $i^*(\lambda)$ of $J(x)$ on $\mathcal{C}(\lambda)$ and $\mathcal{C}^*(\lambda)$ satisfy the relation $i(\lambda) \geq i^*(\lambda)$. If $\mathcal{C}^*$ is the orthogonal complement of a linear subclass $\mathfrak{D}$ of dimension k, then $i(\lambda) \leq i^*(\lambda) + k$, by Theorem 9.5. Hence (16.2) holds.

Comparison theorems for ordinary differential equations of the second order can

be obtained with the help of the following:

THEOREM 16.3. *Let $J^*(x)$ be a second quadratic form on $\mathcal{G}$ such that $J^*(x)$ $\geq J(x)$ on $\mathcal{G}$. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ and $\lambda_1^* \leq \lambda_2^* \leq \cdots \leq \lambda_{m^*}^*$ be respectively the focal points of $J(x)$ and $J^*(x)$ relative to $\mathcal{G}(\lambda)$ ($\lambda' \leq \lambda \leq \lambda''$), each repeated a number of times equal to its order. Then the inequalities*

$$(16.3) \qquad\qquad \lambda_r \leq \lambda_r^* \qquad\qquad\qquad (r = 1, \cdots, m^*)$$

hold. If $J(x) = J^*(x)$ only in case $x = 0$ and $\mathcal{G}(\lambda - 0) = \mathcal{G}(\lambda)$, then

$$(16.4) \qquad\qquad \lambda_r < \lambda_r^* \qquad\qquad\qquad (r = 1, \cdots, m^*).$$

This result follows from Theorem 9.4.

17. An application of the theory of focal points. Consider now the class $\mathcal{B}$ of all arcs

$$x: \qquad\qquad x^j(t) \qquad\qquad\qquad (j = 1, \cdots, q, \ a \leq t \leq b)$$

in the class $\mathcal{G}$ described in §4 vanishing at $t = b$ and satisfying the equations

$$a_{hj} x^j(a) = 0 \qquad\qquad\qquad (h = 1, \cdots, p \leq q),$$

which we assume to be linearly independent. Let $J(x)$ be of the form

$$J(x) = A_{jk} x^j(a) x^k(a) + \int_a^b 2\omega(t, x, \dot{x}) \, dt,$$

where $A_{jk} = A_{kj}$, and 2ω is of the form (4.15). We assume that the strengthened condition of Legendre (4.9) holds. Consequently $J(x)$ is a Legendre form on $\mathcal{G}$, as well as on $\mathcal{B}$.

Let $\mathcal{B}(\lambda)$ ($a \leq \lambda \leq b$) be the set of all arcs x in $\mathcal{B}$ having $x^j(t) = 0$ ($\lambda \leq t \leq b$). It is easily verified that this family has the properties described in the last section provided we replace $\mathcal{G}$ by $\mathcal{B}$. Moreover, the relation $\mathcal{B}(\lambda - 0) = \mathcal{B}(\lambda)$ ($a < \lambda \leq b$) holds.

By an argument like that used in the illustrations given in §6 we can prove the following:

LEMMA 17.1. *Let x be an arc in $\mathcal{G}$ and set*

$$(17.1) \qquad\qquad \xi_j(t) = \int_a^t \omega_{x^j} [s, x(s), \dot{x}(s)] \, ds + c_j.$$

Then the arc x is J-orthogonal to the class $\mathcal{B}(\lambda)$ $(a < \lambda \leq b)$ *if and only if the constants* c_j *in* (17.1) *and constants* e_h *can be chosen such that*

(17.2) $$\xi_j(t) = \omega_{\dot{x}_j}[t, x(t), \dot{x}(t)]$$

holds almost everywhere on $a \leq t \leq \lambda$, *and such that the transversality conditions*

(17.3) $$A_{jk} x^k(a) + e_h a_{hj} = \xi_j(a)$$

hold at $t = a$. *If* $x^j(t) = \xi_j(t) = 0$ *at one point of* $a \leq t \leq \lambda$, *then* $x^j(t) \equiv 0$ *on* $a \leq t \leq \lambda$.

The last statement follows from the theory of differential equations. In view of this result we have the following:

COROLLARY. *No J-transversal of* $\mathcal{B}(\lambda_1)$ *is a J-transversal of* $\mathcal{B}(\lambda_2)$ *if* $\lambda_1 \neq \lambda_2$.

An arc x that is J-orthogonal to $\mathcal{B}$ will be called a *focal arc*. Such an arc is an extremal on $a \leq t \leq b$ satisfying the transversality condition (17.3). We have the following:

THEOREM 17.1. *A value* λ *on* $a < \lambda < b$ *is a focal point of* $J(x)$ *relative to* $\mathcal{B}(\lambda)$ $(a \leq \lambda \leq b)$ *if and only if there is a focal arc* $x \neq 0$ *that vanishes at* $t = \lambda$. *The order of* λ *as a focal point is equal to the number of linearly independent focal arcs of this type in a maximal set.*

This result follows readily from the corollaries to Theorems 16.1 and 17.1.

If $\mathcal{B}^*$ is the class of all arcs in $\mathcal{B}$ having $x^j(a) = 0$, and $\mathcal{B}^*(\lambda)$ is the class of all arcs in $\mathcal{B}^*$ vanishing identically on $\lambda \leq t \leq b$, then the focal points of J relative to $\mathcal{B}^*(\lambda)$ $(a \leq \lambda \leq b)$ are called *conjugate points*. It is easy to see that if $p = 0$ then focal points and conjugate points cannot coincide.

In view of Theorem 16.2 we have:

THEOREM 17.2. *The k-th focal point of* $J(x)$ *relative to* $\mathcal{B}(\lambda)$ $(a \leq \lambda \leq b)$ *precedes (or coincides with) the k-th conjugate point of* $J(x)$, *if the latter exists. Moreover the k-th conjugate point must precede (or coincide with) the* $(k + q - p)$-*th focal point, if the latter exists.*

An analogous theorem can be established with the help of Theorem 16.4. These results serve to illustrate the fact that one can obtain the standard oscillation and comparison theorems in the calculus of variations by use of the results given

in §§15, 16 and 9. These theorems are generalizations of Sturm-Liouville theorems for differential equations.

It should be observed that the results given in §16 also hold for integro-differential equations and integral equations with a symmetric kernel. For example if we replace $J(x)$ by $J(x) + K(x)$, where $K(x) = K(x, x)$ and $K(x, y)$ is given by (4.5), a focal point theory for integro-differential equations is obtained. If $J(x)$ is taken of the form (3.8), and $\mathcal{Q}(\lambda)$ $(a \leq \lambda \leq b)$ is the class of all Lebesgue square integrable functions on $a \leq t \leq b$ having $x(t) = 0$ $(\lambda \leq t \leq b)$, a theory of focal points for integral equations is obtained.

18. Legendre forms depending upon a parameter. Let Λ be a metric space. For each element λ in Λ let $J(x, \lambda)$ be a quadratic form on $\mathcal{Q}$. We shall assume that the quadratic form $J(x; \lambda)$ and the associated bilinear form $J(x, y; \lambda)$ have the following property: Given a sequence of vectors $\{x_r\}$ in $\mathcal{Q}$ converging weakly to a vector x_0 and a sequence $\{\lambda_r\}$ in Λ converging to an element λ_0, then

(a) $\lim\limits_{r=\infty} J(x_r, y; \lambda_r) = J(x_0, y; \lambda_0)$ for every vector y in $\mathcal{Q}$;

(b) $\lim\limits_{r=\infty} \inf J(x_r; \lambda_r) \geq J(x_0; \lambda_0)$;

(c) if $\lim\limits_{r=\infty} J(x_r; \lambda_r) = J(x_0; \lambda_0)$, then $x_r \Longrightarrow x_0$.

It is clear that for each fixed element λ in C the form $J(x; \lambda)$ is a Legendre form on $\mathcal{Q}$. These properties are enjoyed by

$$J(x; \lambda) = J(x) - \lambda K(x),$$

where $J(x)$ is a Legendre form, $K(x)$ is a w-continuous form, and λ is a real number. As a second example, one should observe that the second variation $J(x, \lambda)$ of a function $F(\lambda)$ in the calculus of variations will have this property in a weak neighborhood of an arc λ satisfying the strengthened condition of Legendre.

LEMMA 18.1. *Let $\mathcal{B}$ be a closed subclass of $\mathcal{Q}$ on which $J(x; \lambda_0)$ is positive, where λ_0 is in Λ. Then there is a neighborhood N of λ_0 and a positive number h such that the inequality*

(18.1) $$J(x; \lambda) \geq h\,|x|^2$$

holds on $\mathcal{B}$ for every element λ in N.

Suppose the conclusion in the lemma is false. Then there is a sequence of unit

vectors $\{x_r\}$ in $\mathcal{B}$ converging weakly to a vector x_0, and a sequence of elements λ_r in Λ converging to λ_0, such that

$$J(x_r; \lambda_r) < 1/r .$$

Clearly x_0 is in $\mathcal{B}$ since $\mathcal{B}$ is closed. Moreover

$$0 \geq \lim_{r=\infty} \sup J(x_r; \lambda_r) \geq \lim_{r=\infty} \inf J(x_r; \lambda_r) \geq J(x_0; \lambda_0) \geq 0 .$$

Hence $J(x_0; \lambda_0) = 0$ and $x_0 = 0$. Consequently

$$\lim_{r=\infty} J(x_r; \lambda_r) = J(x_0; \lambda_0) = 0 .$$

By virtue of our hypotheses regarding $J(x; \lambda)$, we have $x_r \Longrightarrow x_0 = 0$, contrary to the relation $|x_r| = 1$. This proves the lemma.

LEMMA 18.2. *Let* $\{x_r\}$ *be a sequence of unit vectors converging weakly to a vector* x_0, *and let* $\{\lambda_r\}$ *be a sequence of elements in* Λ *converging to an element* λ_0. *If for each positive integer r the relation*

(18.2) $J(x_r, y; \lambda_r) = 0$

holds for every y in $\mathcal{Q}$, *then this relation holds for r = 0. Moreover* $|x_0| = 1$ *and* $x_r \Longrightarrow x_0$.

The first conclusion follows from property (a) of $J(x; \lambda)$. By (18.2), we have $J(x_r; \lambda_r) = J(x_0; \lambda_0) = 0$. Consequently $x_r \Longrightarrow x_0$, by property (c). Since the vectors x_r are unit vectors, so also is x_0.

THEOREM 18.1. *Let* $i(\lambda)$ *be the index and* $n(\lambda)$ *be the nullity of* $J(x; \lambda)$ *on* $\mathcal{Q}$ *for each element* λ *in* Λ. *Given an element* λ_0 *in* Λ, *there is a neighborhood N of* λ_0 *in which the inequalities*

(18.3) $i(\lambda_0) \leq i(\lambda) \leq i(\lambda) + n(\lambda) \leq i(\lambda_0) + n(\lambda_0)$

hold. If $n(\lambda_0) = 0$, *then* $n(\lambda) = 0$, $i(\lambda) = i(\lambda_0)$ *on N.*

Let $\mathcal{B}$ be a maximal linear subset of $\mathcal{Q}$ on which $J(x; \lambda_0)$ is negative. Its dimension is $i(\lambda_0)$. By continuity considerations it is seen that there is a neighborhood N of λ_0 such that $J(x; \lambda)$ is negative on $\mathcal{B}$ whenever λ is in N. Consequently $i(\lambda) \geq i(\lambda_0)$ on N, and the first inequality in (18.3) is established.

In order to prove the last inequality in (18.3), observe that by virtue of Theorem 9.2, the form $J(x; \lambda_0)$ is positive on the orthogonal complement $\mathfrak{B}$ of a linear subclass $\mathcal{C}$ of dimension $k = i(\lambda_0) + n(\lambda_0)$. By Lemma 18.1, the neighborhood N chosen above can be diminished so that $J(x; \lambda)$ is positive on $\mathfrak{B}$ whenever λ is in N. Using Theorem 9.2 again, we see that $i(\lambda) + n(\lambda) \leq k$, as was to be proved.

COROLLARY 1. *The set of points λ in Λ at which $n(\lambda) \neq 0$ is a closed subset of Λ.*

COROLLARY 2. *Given an element λ_0 in Λ, there is a neighborhood N of λ_0 on which the inequality $n(\lambda) \leq n(\lambda_0)$ holds.*

As an application of Theorem 18.1 to the calculus of variations one obtains the following result: If an extremal λ_0 satisfying the strengthened condition of Legendre is nondegenerate and of index i, then neighboring extremals are nondegenerate and of index i.

THEOREM 18.2. *Suppose that Λ is an open interval of real numbers. If for a value λ_0 in Λ there is a number $\epsilon > 0$ such that given two values $\lambda_1 < \lambda_2$ in the ϵ-neighborhood of λ_0 the relation $J(x; \lambda_2) < 0$ holds whenever $x \neq 0$ and $J(x; \lambda_1) \leq 0$, then ϵ can be chosen such that*

$$(18.4) \qquad i(\lambda) = i(\lambda_0), \quad n(\lambda) = 0 \qquad (\lambda_0 - \epsilon < \lambda < \lambda_0),$$

$$i(\lambda) = i(\lambda_0) + n(\lambda_0), \quad n(\lambda) = 0 \qquad (\lambda_0 < \lambda < \lambda_0 + \epsilon).$$

Similarly, if for a value λ_0 in Λ there is a number $\epsilon > 0$ such that given two values $\lambda_1 < \lambda_2$ in the ϵ-neighborhood of λ_0 the relation $J(x, \lambda_1) < 0$ holds whenever $x \neq 0$ and $J(x; \lambda_2) \leq 0$, then ϵ can be chosen so that

$$(18.5) \qquad i(\lambda) = i(\lambda_0) + n(\lambda_0), \quad n(\lambda) = 0 \qquad (\lambda_0 - \epsilon < \lambda < \lambda_0),$$

$$i(\lambda) = i(\lambda_0), \quad n(\lambda) = 0 \qquad (\lambda_0 < \lambda < \lambda_0 + \epsilon).$$

In view of Theorems 9.1 and 9.2, the relation $i(\lambda_2) \geq i(\lambda_1) + n(\lambda_1)$ will hold if $J(x, \lambda_2) < 0$ whenever $x \neq 0$ and $J(x, \lambda_1) \leq 0$. Combining this result with (18.3), we find that (18.4) holds provided ϵ is sufficiently small. The last statement in the theorem follows similarly.

COROLLARY 1. *If $J(x; \lambda_1) > J(x; \lambda_2)$ whenever $x \neq 0$, and λ_1, λ_2 are in Λ with $\lambda_1 < \lambda_2$, then the relations (18.4) hold for all λ_0 in Λ.*

A value λ will be called a *characteristic value* if $n(\lambda) \neq 0$; and $n(\lambda)$ will be called its *multiplicity*. A characteristic value will be counted a number of times equal to its multiplicity. If λ is a characteristic number, then a vector $x \neq 0$ such that $J(x, y, \lambda) = 0$ for all y will be called a *characteristic vector* corresponding to λ.

COROLLARY 2. *If* $J(x; \lambda_1) > J(x, \lambda_2)$ *whenever* $x \neq 0$ *and* $\lambda_1 < \lambda_2$, *then for* $\lambda' < \lambda''$ *the difference* $i(\lambda'') - i(\lambda')$ *is equal to the number of characteristic values on* $\lambda' \leq \lambda < \lambda''$. *If there is a value* λ^* *at which* $i(\lambda^*) = 0$, *then* $i(\lambda)$ *is equal to the number of characteristic values less than* λ.

As a further result we have:

THEOREM 18.3. *Let* $J(x)$ *be a Legendre form on* $\mathcal{U}$ *and let* $K(x)$ *be a w-continuous quadratic form such that* $J(x) > 0$ *whenever* $x \neq 0$ *and* $K(x) \leq 0$. *If*

$$(18.6) \qquad J(x; \lambda) = J(x) - \lambda K(x),$$

then there is a value λ^* *such that* $J(x; \lambda^*)$ *is positive definite on* $\mathcal{U}$. *If* $\lambda_0 \geq \lambda^*$, *the relations* (18.4) *hold; and if* $\lambda_0 \leq \lambda^*$, *the relations* (18.5) *hold. If* $\lambda^* \leq \lambda' < \lambda''$, *then* $i(\lambda'') - i(\lambda')$ *is equal to the number of characteristic values on* $\lambda' \leq \lambda < \lambda''$; *and if* $\lambda'' < \lambda' \leq \lambda^*$ *then* $i(\lambda'') - i(\lambda')$ *is equal to the number of characteristic values on* $\lambda'' < \lambda \leq \lambda'$. *The index* $i(\lambda)$ *of* $J(x; \lambda)$ *on* $\mathcal{U}$ *is equal to the number of characteristic values between* λ *and* λ^*.

The existence of the number λ^* follows from Theorem 13.2. From the relation

$$J(x; \lambda) = J(x, \lambda^*) + (\lambda^* - \lambda) K(x),$$

we conclude that if $\lambda < \lambda^*$, $J(x, \lambda) \leq 0$, and $x \neq 0$, then $K(x) < 0$. Consequently, if $\lambda_1 < \lambda_2 < \lambda^*$ we have

$$J(x; \lambda_1) = J(x; \lambda_2) + (\lambda_2 - \lambda_1) K(x) < 0$$

whenever $x \neq 0$ and $J(x, \lambda_2) \leq 0$. The relations (18.5) therefore hold when $\lambda_0 \leq \lambda^*$. Similarly (18.4) hold when $\lambda_0 \geq \lambda^*$. The remaining statements in the theorem follow readily.

COROLLARY 1. *If* $K(x) \geq 0$ *on* $\mathcal{U}$, *then the index* i *of* $J(x)$ *on* $\mathcal{U}$ *is equal to the number of negative characteristic values.*

In view of the connection between $i(\lambda)$ and the number of characteristic values

we have:

COROLLARY 2. *If $K(x)$ is nondegenerate on a linear subclass $\mathcal{B}$ of $\mathcal{Q}$ of dimension h, then there are at least h characteristic numbers. If $K(x)$ is nondegenerate on $\mathcal{Q}$, then there is a denumberable set of characteristic numbers.*

The following result is easily established.

THEOREM 18.4. *Let $J(x, \lambda)$ be of the form (18.6). If x' and x'' are characteristic vectors corresponding to distinct characteristic values λ' and λ'', then*

$$J(x', x'') = K(x', x'') = 0 \,.$$

Corresponding to each characteristic value λ, there are $r = n(\lambda)$ linearly independent characteristic vectors $x_1, \cdots, x_r$ corresponding to λ which can be chosen so that

$$(18.7) \qquad K(x_p) = k, \;\; J(x_p) = \lambda k, \;\; J(x_p, x_q) = K(x_p, x_q) = 0$$

$$(p \neq q, p, q = 1, \cdots, r) \,,$$

where $k = -1$ or $k = 1$ according as $\lambda < \lambda^*$ or $\lambda > \lambda^*$. Every characteristic vector corresponding to λ is expressible in the form $x = x_p a_p$.

In view of this result, the sequence $\{\lambda_p\}$ of characteristic values, each repeated a number of times equal to its multiplicity, has associated with it a sequence of characteristic vectors $\{x_p\}$ such that

$$(18.8) \qquad J(x_p, x_q) = j_p \delta_{pq}, \;\; K(x_p, x_q) = k_p \delta_{pq} \qquad (j_p = \lambda_p k_p, \;\; k_p = \pm 1),$$

where $\delta_{pp} = 1$, $\delta_{pq} = 0$ $(p \neq q)$. Only a finite number of the j's are negative. We can accordingly suppose that

$$j_1 \leq j_2 \leq \cdots \leq j_p \leq \cdots; \quad \lambda_p \leq \lambda_q \qquad \text{if } j_p = j_q \text{ and } p < q \,.$$

The following well-known result can be established by the use of the theory of indices developed in the preceding pages. Its proof will be omitted.

THEOREM 18.5. *Let $J(x, \lambda)$ be of the form (18.6). Let a vector x in $\mathcal{Q}$ be given, and set*

$$a_q = K(x_q, x)/k_q \qquad\qquad (q = 1, 2, 3, \cdots) \,.$$

Then the sequence $\{y_q\}$ defined by the formula

$$y_q = a_1 x_1 + \cdots + a_{_{ij}} x_q$$

converges to a vector y in $\mathfrak{A}$; moreover, $J(y)$ and $K(y)$ are given by the absolutely convergent series

$$J(y) = j_1 a_1^2 + j_2 a_2^2 + \cdots + j_q a_q^2 + \cdots,$$

$$K(y) = k_1 a_1^2 + k_2 a_2^2 + \cdots + k_q a_q^2 + \cdots;$$

the vector $z = x - y$ is a K-transversal of $\mathfrak{A}$ and

$$J(x) = J(y) + J(z), \quad K(x) = K(y), \quad K(z) = 0 ;$$

if $K(x)$ is nondegenerate on $\mathfrak{A}$, then $x = y$.

19. Further comparison theorems. In the present section we assume that $J(x; \lambda)$ is of the form (18.6) and that $K(x)$ is positive on $\mathfrak{A}$. Then, in (18.8), we have $j_p = \lambda_p$ and $\lambda_1 \leq \lambda_2 \leq \cdots$. As before, $i(\lambda)$ denotes the index and $n(\lambda)$ the nullity of $J(x, \lambda)$ on $\mathfrak{A}$.

LEMMA 19.1. *Given a number λ, the number of characteristic values $\leq \lambda$ is equal to the dimension $i(\lambda) + n(\lambda)$ of a maximal linear subclass of $\mathfrak{A}$ on which $J(x; \lambda) \leq 0$. The number of characteristic values less than λ is given by the dimension $i(\lambda)$ of a maximal linear subclass of $\mathfrak{A}$ on which $J(x; \lambda)$ is negative.*

This result follows from Theorem 18.3. Using this result we can prove the minimax property of characteristic values stated in the following:

THEOREM 19.1. *Let C be a linear subclass of $\mathfrak{A}$ of dimension r, and let λ^* be the maximum of the quotient $J(x)/K(x)$ for all $x \neq 0$ in C. Then $\lambda^* \geq \lambda_r$, the r-th characteristic value of $J(x, \lambda)$ on $\mathfrak{A}$. Moreover there is a subclass C of $\mathfrak{A}$ of dimension r on which λ_r is the maximum of $J(x)/K(x)$.*

For clearly $J(x, \lambda^*) \leq 0$ on C. By Lemma 19.1, we have $r \leq i(\lambda^*) + n(\lambda^*)$, the number of characteristic numbers less than or equal to λ^*. Hence $\lambda_r \leq \lambda^*$. The linear subclass C of $\mathfrak{A}$ generated by the characteristic vectors $x_1, \cdots, x_r$ corresponding to $\lambda_1, \cdots, \lambda_r$ is such that $\lambda_r = J(x_r)/K(x_r)$ is the maximum of $J(x)/K(x)$ on C. Hence $\lambda^* = \lambda_r$ in this case.

THEOREM 19.2. *Let $\mathfrak{A}^*$ be a closed linear subclass of $\mathfrak{A}$ and let $\lambda_1^* \leq \lambda_2^* \leq \cdots \leq \lambda_r^*$ be the characteristic values of $J(x, \lambda)$ on $\mathfrak{A}^*$, each repeated a number of times equal to its multiplicity. The r-th characteristic value λ_r^* of $J(x, \lambda)$ on*

$\mathcal{C}^*$, *if it exists, is greater than or equal to the r-th characteristic value* λ_r *of* $J(x, \lambda)$ *on* $\mathcal{C}$.

For by the last theorem there is a linear subclass $\mathcal{C}$ of $\mathcal{C}^*$ of dimension r such that λ_r^* is the maximum of $J(x)/K(x)$ on $\mathcal{C}$. Hence $\lambda_r^* \geq \lambda_r$, again by Theorem 19.1.

THEOREM 19.3. *Let* $J^*(x)$ *be a second Legendre form on* $\mathcal{C}$ *such that* $J(x) \leq J^*(x)$ *on* $\mathcal{C}$*, and set*

$$J^*(x, \lambda) = J^*(x) = \lambda K(x).$$

Denote by $\lambda_1^* \leq \lambda_2^* \leq \cdots \leq \lambda_r^* \leq \cdots$ *the characteristic values of* $J^*(x, \lambda)$ *on* $\mathcal{C}$*, each repeated a number of times equal to its multiplicity. Then* $\lambda_r \leq \lambda_r^*$ $(r = 1, 2, 3, \cdots)$. *Moreover if* $J(x) < J^*(x)$ *whenever* $x \neq 0$*, then* $\lambda_r < \lambda_r^*$ $(r = 1, 2, 3, \cdots)$.

Let $\mathcal{C}$ be a linear subclass of $\mathcal{C}$ of dimension r on which $J^*(x, \lambda_r^*) \leq 0$. Then also $J(x, \lambda_r^*) \leq 0$ on $\mathcal{C}$. Hence $\lambda_r^* \geq \lambda_r$, by Theorem 19.1. If $J^*(x) > J(x)$ when $x \neq 0$, then $J(x, \lambda_r^*)$ is negative on $\mathcal{C}$. It follows that $\lambda_r < \lambda_r^*$, again by Theorem 19.1.

REFERENCES

1. A. A. Albert, *A quadratic form problem in the calculus of variations*, Bull. Amer. Math. Soc. 44 (1938), 250-252.

2. G. D. Birkhoff and M. R. Hestenes, *Natural isoperimetric conditions in the calculus of variations*, Duke Math. J. 1 (1935), 198-286.

3. G. A. Bliss, *A boundary value problem for a system of ordinary differential equations of the first order*, Trans. Amer. Math. Soc. 28 (1926), 561-584.

4. ———, *Lectures on the calculus of variations*, The University of Chicago Press, 1946.

5. J. W. Calkin, *Functions of several variables and absolute continuity I*, Duke Math. J. 6 (1940), 170-186.

6. L. M. Graves, *The Weierstrass condition for multiple integral variation problems*, Duke Math. J. 5 (1939), 656-660.

7. K. Hazard, *Index theorems for the problem of Bolza in the calculus of variations*, Contributions to the calculus of variations, 1938-1941, The University of Chicago Press, pp. 293-356.

8. M. R. Hestenes, *Sufficient conditions for multiple integral problems in the calculus of variations*, Amer. J. Math. 70 (1948), 239-275.

9. M. R. Hestenes and E. J. McShane, *A theorem on quadratic forms and its application in the calculus of variations*, Trans. Amer. Math. Soc. 47 (1940), 501-512.

10. W. Karush, *Isoperimetric problems and index theorems in the calculus of variations*, Dissertation, The University of Chicago, 1942.

11. C. B. Morrey, Jr., *Functions of several variables and absolute continuity II*, Duke Math. J. 6 (1940), 187-215.

12. ———, *Multiple integral problems in the calculus of variations and related topics*, University of California Publications in Mathematics, New Series, 1 (1943), 1-130.

13. M. Morse, *The calculus of variations in the large*, American Mathematical Society Colloquium Publications, vol. 18, New York, 1934.

14. W. T. Reid, *Boundary value problems of the calculus of variations*, Bull. Amer. Math. Soc. 43 (1937), 633-666.

15. ———, *A new class of self-adjoint boundary value problems*, Trans. Amer. Math. Soc. 52 (1942), 381-425.

16. ———, *A theorem on quadratic forms*, Bull. Amer. Math. Soc. 44 (1938), 437-440.

17. L. Ritcey, *Index theorems for discontinuous problems in the calculus of variations*, Dissertation, The University of Chicago, 1945.

18. M. H. Stone, *Transformations in Hilbert space and their applications to analysis*, American Mathematical Society Colloquium Publications, vol. 15, New York, 1932.

19. Béla v. Sz.-Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 5, part 5, Springer, Berlin, 1942.

20. F. J. Terpstra, *Die Darstellung biquadratischer Formen als Summen von Quadraten mit Anwendung auf die Variationsrechnung*, Math. Ann. 116 (1939), 166-180.

UNIVERSITY OF CALIFORNIA, LOS ANGELES, AND
NATIONAL BUREAU OF STANDARDS, LOS ANGELES