A Generalized Sturm Theorem*

By H. M. EDWARDS**

Introduction

This paper contains a generalization of the Sturm oscillation, comparison, and separation theorems to the case of formally self-adjoint linear systems of even order. With the aid of a reformulation of the theory in terms of the calculus of variations, we are able to give an elementary and self-contained development in which the higher order cases are almost as simple as the usual second order case. The proofs are greatly simplified by the introduction of a new analytic tool (U-manifolds) which is discussed later in this introduction.

The Sturm theorems deal with differential equations of the type

$$-(px')' + rx = \lambda x \tag{0.1}$$

where $'$ denotes differentiation, and where p and r are given (differentiable) functions with $p > 0$. If we set $\Omega[x, y] = px'y' + rxy$, then integration by parts shows that $x(t)$ is a solution of (0.1) for $t \in [a, b]$ if and only if $\int_a^b (\Omega[x, y] - \lambda xy)dt$ is zero for all $y(t)$ satisfying $y(a) = y(b) = 0$.

Let $V[a, b]$ be the vector space of all (differentiable) functions x on $[a, b]$, and let $V_0[a, b]$ be the subspace consisting of those x for which $x(a) = x(b) = 0$. Let $\Omega[x]$ be the quadratic expression $\Omega[x, x] = px'^2 + rx^2$. Recalling that the *nullity* of a quadratic form is the degree of degeneracy of the corresponding symmetric bilinear form, the above shows that: the number of linearly independent solutions x of (0.1) satisfying $x(a) = x(b) = 0$ is equal to the nullity of the quadratic form $x \to \int_a^b (\Omega[x] - \lambda x^2)dt$ on $V_0[a, b]$. Then the Sturm oscillation theorem (number of zeros = number of negative eigenvalues) can be restated as:

$$\begin{aligned}\sum_{a<t<b} \left\{\text{nullity of } x \to \int_a^t \Omega[x]dt \text{ on } V_0[a, t]\right\} \\ = \sum_{\lambda<0} \left\{\text{nullity of } x \to \int_a^b (\Omega[x] - \lambda x^2)dt \text{ on } V_0[a, b]\right\}.\end{aligned} \tag{0.2}$$

This has little to recommend it over the usual statement. The objection to the usual statement is that it obscures the third and most concise description of the number in question, namely that it is:

* This work has been partially supported by the Office of Naval Research under Contract NONR 266 (57).

** The author holds a National Science Foundation Post-Doctoral Fellowship.

$$\text{(0.3)} \qquad \text{the index of } x \to \int_a^b \Omega[x]dt \text{ on } V_0[a, b] .$$

(The *index* of a quadratic form is the maximum possible dimension of a subspace on which it is negative definite.) The theorem which is generalized in this paper is:

STURM THEOREM. LHS(0.2) = RHS(0.2) = (0.3).

It is generalized in the following ways:

(a) Ω may depend on any number of derivatives provided it satisfies a condition analogous to "$p > 0$."

(b) $V[a, b]$ is generalized to be complex vector-valued functions on $[a, b]$.

(c) The domain of the quadratic forms may be extended from $V_0[a, b]$ to $V[a, b]$ and they may take the form $x \to \left(\int_a^b \Omega[x]\right) + \beta[x]$ where β is a quadratic form depending on the values and derivatives of x at a and b.

(d) RHS(0.2) is generalized to $\sum_{\lambda<0}\left\{\text{nullity of } \int_a^b (\Omega - \lambda\Lambda)dt\right\}$ where Λ, like Ω, may depend on derivatives, provided that $\int_a^b \Lambda$ is positive definite on $V[a, b]$.

Generalizations (a)–(c) have the effect of generalizing the Sturm theorem from the case of (0.1) subject to $x(a) = x(b) = 0$ to the case of an arbitrary formally self-adjoint system (real or complex) of even order subject to an arbitrary self-adjoint boundary condition (§ 6). Generalization (d) is included only to strengthen the following analogy: Let E be a finite dimensional vector space, and let Q_1, Q_2 be quadratic forms on E with Q_2 positive definite. Then

$$\text{(0.4)} \qquad \text{index of } Q_1 = \sum_{\lambda>0} \text{nullity } (Q_1 + \lambda Q_2)$$

as is seen by writing Q_1 in diagonal form relative to the norm Q_2; i.e. by using the spectral theorem (finite dimensional case). The equality (0.3) = RHS(0.2) is the statement that (0.4) holds for Q_1, Q_2 in the class of quadratic forms described by (a)–(c).

Consider now the equality (0.2). The simplest proof of it runs as follows (cf. Levinson and Coddington [3, pp. 209–212]): Let C be the circle consisting of all lines through the origin of the euclidean plane (i.e., C is the real projective line) and define $c(\lambda, t)$ = the line containing $(x_\lambda(t), x'_\lambda(t))$ where x is any non-trivial solution of (0.1) with $x_\lambda(a) = 0$. Let $p \in C$ be the line through (0, 1). Then $c(\lambda, a) \equiv p$, and for $t > a$, $c(\lambda, t) = p$ if and only if there is a non-trivial solution of (0.1) satisfying $x(a) = x(t) = 0$. Thus (0.2) is equivalent to

$$\text{(0.5)} \qquad \begin{aligned} &\{\text{number of times that } c(0, t) = p \text{ for } a < t < b\} \\ &\quad = \{\text{number of times that } c(\lambda, b) = p \text{ for } \lambda < 0\} . \end{aligned}$$

To prove this, one first proves the following statements:

(1) If λ is fixed and t increases, then $c(\lambda, t)$ moves in a monotone direction around C. [Orient C by calling this direction "clockwise".]

(2) If $t > a$ is fixed and λ decreases, then $c(\lambda, t)$ moves monotone counterclockwise.

(3) There is a λ^- such that $c(\lambda, t) \neq p$ whenever $t \in (a, b]$, $\lambda \leqq \lambda^-$.

Then (0.5) is seen as follows: Consider c restricted to the boundary of the rectangle $[\lambda^-, 0] \times [a, b]$. On the side $t = a$ it is constantly p. On the side $\lambda = \lambda^-$ it is p only for $t = a$. From the fact that the total number of times that the curve winds around C is zero (because c is defined on the whole rectangle) the intersections with p on the other two sides must cancel, hence (0.5).

This is precisely the proof that we give for the generalization of (0.2) (I of the proof of Theorem 3.1). In the general case C becomes a "U-manifold", "clockwise curves" become "$\oplus$-curves", and the intuitive geometric argument by which (0.5) is deduced from (1)–(3) becomes an argument based on a certain theory of multiplicities of intersections of curves in U-manifolds with certain subvarieties Γ of codimension 1 [in (0.5), $\Gamma = p$]. These generalizations are presented separately (in § 4) but should be read concurrently with §§ 1–3.

The primary emphasis of the presentation is on the formulation of the problem and on the method of proof. With this end in mind, the exposition has been made almost entirely self-contained, which has had the disadvantage of blurring all distinctions between new work and classical results. Generally speaking, the results are not new. For example, the main theorem (Theorem 3.1) can be deduced from a strong version of the spectral theorem (in particular a version which includes the regularity, i.e. differentiability or analyticity of the eigenfunctions, and discreteness of the spectrum); but, as is shown in § 5, the implication can be reversed to prove such a spectral theorem for a large class of ordinary differential operators and boundary conditions. Such spectral theorems are normally called Sturm-Liouville theorems, at least in the case of second order operators. In this case, which arises in the classical calculus of variations, Theorem 3.1 is due to Morse [5] who was the first to generalize the Sturm theorems to the case of vector-valued operators. Many of the methods are not original with this paper either. For example the intersection-theoretic method is due to Bott [2] in the second order case. The fundamental importance of considering the form Ω, rather than the differential operator L, seems to be one of those facts which date back to the beginnings of the subject but which nonetheless are constantly being rediscovered. Finally, I would like to acknowledge my indebtedness to the papers of Ambrose [1] and Heinz [4] which deal with theorems not treated here, but which suggested many of the methods used.

The contents of the paper are as follows: § 1 defines the basic notions of "derivative dependent hermitian form", "Sturm form", and "solution space", and proves a theorem about the extent to which the solution space determines the form (Theorem 1.2). Although this theorem is extraneous to what follows, it is nonetheless interesting in itself, and its proof is important in § 2. The notion of "index problem" is defined in § 2. Theorem 2.1 shows that for such problems only Sturm forms are of interest. Propositions 2.4 and 2.6 provide important lemmas in the proof of the main theorem. § 3 is devoted to the statement and proof of Theorem 3.1. § 4 contains the theory of "U-manifolds". § 5 relates Theorem 3.1 to the theory of eigenfunction expansions. § 6 relates "index problems" to the classical boundary value problems consisting of a formally self-adjoint linear differential operator L of even order with positive definite leading coefficient together with a self-adjoint boundary condition. § 7 gives the generalizations of the Sturm-Morse comparison and separation theorems which result from Theorem 3.1. § 8 includes an elementary exposition of the calculus of variations and shows the role of Sturm forms in this theory.

I would like to take advantage of my first opportunity to express publicly my gratitude to Prof. Raoul Bott for the honor of having been his student. Those who are familiar with his paper [2] will recognize the extent of his influence on the subject matter here; I would be pleased to think that it bears his mark in other ways as well.

1. Derivative dependent quadratic forms

1.1. *Notation.* A (real-valued) *hermitian form* on a complex vector space V is a real-valued function Q on V which satisfies:

(a) The parallelogram law: $Q[v_1 + v_2] + Q[v_1 - v_2] = 2(Q[v_1] + Q[v_2])$ all $v_1, v_2 \in V$,

(b) $Q[cv] = |c|^2 Q[v]$ all $c \in \mathbf{C}$, $v \in V$.

Such a function is of the form $Q[v] = Q[v, v]$ where $Q: V \times V \to \mathbf{C}$ is a uniquely determined hermitian symmetric sesqui-linear form. (Following Bourbaki, *sesqui-linear* means linear in the first variable, conjugate linear in the second.) In the case where V is finite-dimensional a hermitian form can also be described as follows:

For a finite-dimensional complex vector space E we define E^* to be the space of all linear maps $E \to \mathbf{C}$ with addition and scalar multiplication defined in such a way that the natural map $E \times E^* \to \mathbf{C}$ is (linear in the first variable and) *conjugate* linear in the second. This map will be denoted by brackets $\langle\, ,\rangle$. Then any map $\alpha\colon E \rightarrow E^*$ has a dual $\alpha^*\colon (E^*)^* \rightarrow E^*$. On the other hand, there is a natural identification $E \approx (E^*)^*$ and $\alpha = \alpha^*$ if and only if the map $v \rightarrow \langle v, \alpha v\rangle$ of $E \rightarrow \mathbf{C}$ is real-valued. Then the rule $\alpha \leftrightarrow \langle v, \alpha v\rangle$ is a one-one correspondence between self-dual maps $E \rightarrow E^*$ and hermitian forms on E.

According to context, a hermitian form will be interpreted in any of these three ways. In particular, if Q is a hermitian form, then $Q[v_1, v_2]$ will denote the value of the corresponding hermitian symmetric sesqui-linear form.

Associated with a hermitian form Q on V we have the following notions: Q is *positive definite* (resp. *positive semi-definite, negative definite, negative semi-definite*) if $v \neq 0$ implies $Q[v] > 0$ (resp. $\geqq 0$, < 0, $\leqq 0$). Elements v_1, v_2 are *orthogonal relative to* Q if $Q[v_1, v_2] = 0$. The *orthogonal complement relative to* Q of a subspace V_1 will be denoted $V_1^{\perp}$ (rel. Q). Q is *non-degenerate* if $V^{\perp}$ (rel. Q) $= \{0\}$. A decomposition of V as a direct sum of subspaces $V = V_1 \oplus V_2$ (in the algebraic sense) *splits* Q if $Q[v_1 + v_2] = Q[v_1] + Q[v_2]$ for all $v_1 \in V_1$, $v_2 \in V_2$, and this is true if and only if $V_2 \subset V_1^{\perp}$ (rel. Q) or equivalently $V_1 \subset V_2^{\perp}$ (rel. Q).

The notation $V[a, b]$ where $a < b$ are real numbers will be used to denote the complex vector space of all C^r curves $x\colon [a, b] \rightarrow E$ mapping the closed interval into a finite-dimensional complex vector space E; that is, the degree of differentiability r and the image space E will be implicit. The value of r can be a finite integer (r continuous derivatives), ∞ (infinitely differentiable) or ω (analytic). Otherwise stated, $V[a, b]$ is the space of all cross-sections of $E \times [a, b] \rightarrow [a, b]$ considered as a complex vector bundle of class C^r. We have in mind the application of § 8 in which $V[a, b]$ is simply the space of all cross-sections of a complex vector bundle $B \rightarrow [a, b]$ with no natural trivialization $B \approx E \times [a, b]$. For the sake of simplicity we always consider $V[a, b]$ relative to a fixed trivialization as above, but all notions considered will be independent of the trivialization.

For $x \in V[a, b]$, $t \in [a, b]$, and ν an integer $\leqq r + 1$, we denote $(x(t), x'(t), x''(t), \cdots, x^{(\nu-1)}(t)) \in E^{\nu}$, where $'$ denotes derivative, by $\mathbf{x}^{(\nu)}(t)$, dropping the superscript whenever its value is clear from the context. (As an element of E^{ν}, $\mathbf{x}^{(\nu)}(t)$ depends on the choice of trivialization. However, defining $O^{(\nu)}(t) \subset V[a, b]$ by $O^{(\nu)}(t) = \{x \in V[a, b] : \mathbf{x}^{(\nu)}(t) = 0\}$, $O^{(\nu)}(t)$ is independent of the choice of trivialization and $\mathbf{x}^{(\nu)}(t)$ represents the class of x in $V[a, b]/O^{(\nu)}(t)$.)

DEFINITION. A *derivative dependent hermitian form of order* $\leqq \nu$ on $V[a, b]$ is a rule Ω which assigns to each $t \in [a, b]$ a hermitian form $\Omega(t)$ on V

such that:

(a) In the case where r is finite, we assume that $2\nu \leqq r$.

(b) $\Omega(t)[x]$ depends only on $\mathbf{x}^{(\nu+1)}(t)$; i.e., only on the class of x in $V[a, b]/O^{(\nu+1)}(t)$.

(c) $t \to \Omega(t)[x]$ is of class $C^{r-\nu}$ for all $x \in V[a, b]$.

REMARKS. 1. (b) and (c) are not meaningful unless $\nu \leqq r$. The reason for (a) is seen in the proof of Proposition 1.1 below.

2. Note that (c) requires as much differentiability as possible if Ω is to depend non-trivially on $\mathbf{x}^{(\nu+1)}(t)$, which is $C^{r-\nu}$.

Such an Ω can be uniquely written as

$$\Omega(t)[x] = \textstyle\sum_{i,j=0}^{\nu} \langle x^{(i)}(t), \omega_{ij}(t)x^{(j)}(t)\rangle \tag{1.1}$$

where the $\omega_{ij}(t)\colon E \to E^*$ satisfy $\omega_{ji} = \omega_{ij}^*$ and are $C^{r-\nu}$ in t. The simplest case is that in which $\omega_{\nu\nu}(t)$ is non-singular for all t. Invariantly this can be stated: Ω will be called *non-degenerate of order* ν if, for each t, it gives a non-degenerate hermitian form on $O^{(\nu)}(t)/O^{(\nu+1)}(t) \approx E$. This form will be called the *leading coefficient* of Ω (at t). Ω will be called a *Sturm form* if its leading coefficient is positive definite (for all t).

This paper is concerned with hermitian forms on $V[a, b]$ of the type $x \to \int_a^b \Omega(t)[x]dt$ where Ω is a Sturm form on $V[a, b]$ (although for the remainder of this paragraph we merely require that Ω be non-degenerate). "Ω as a hermitian form on $V[a, b]$" will always mean the integral, and integrals will be written without the dt, since at all times there is only one variable, that of the interval $[a, b]$.

1.2 *Solutions.*

DEFINITION. Let Ω be a derivative dependent hermitian form on $V[a, b]$. $x \in V[a, b]$ will be called a *solution* of Ω if it is orthogonal rel. Ω to $O^{(n)}(a) \cap O^{(n)}(b)$ for some integer $n \leqq r$.

PROPOSITION 1.1. *Let Ω be non-degenerate of order ν, and let S be the set of all solutions of Ω. Then:*

(a) *$S \subset V[a, b]$ is a subspace of dimension $2\nu \cdot \dim E$.*

(b) *For each $t \in [a, b]$ we have $S \cap O^{(2\nu)}(t) = \{0\}$, hence the map $S \to E^{2\nu}$ given by $x \to \mathbf{x}^{(2\nu)}(t)$ is 1–1 onto for all t.*

(c) *The elements of S are in fact orthogonal to $O^{(\nu)}(a) \cap O^{(\nu)}(b)$* rel. Ω.

(d) *For any $[t_0, t_1] \subset [a, b]$ the solutions of the restriction of Ω to $V[t_0, t_1]$ are the restrictions of the elements of S.*

PROOF. Rewriting (1.1) as a hermitian sesqui-linear form and integrating by parts formally we have

$$
\begin{aligned}
\Omega(t)[x, y] &= \textstyle\sum_{i,j=0}^{\nu} \langle x^{(i)}(t), \omega_{ij}(t)y^{(j)}(t)\rangle \\
&= \textstyle\sum_{i=0}^{2\nu} \langle x(t), p_i(t)y^{(i)}(t)\rangle + \dfrac{d}{dt}\textstyle\sum_{j=0}^{\nu-1}\sum_{k=0}^{2\nu-j-1} \langle x^{(j)}(t), a_{jk}(t)y^{(k)}(t)\rangle \qquad (1.2) \\
&= \langle x(t), L(t)\mathbf{y}^{(2\nu+1)}(t)\rangle + \dfrac{d}{dt}\langle \mathbf{x}^{(\nu)}(t), A(t)\mathbf{y}^{(2\nu)}(t)\rangle
\end{aligned}
$$

where $L(t)$: $E^{2\nu+1} \to E^*$ and $A(t)$: $E^{2\nu} \to E^{\nu *}$ are so defined. The only specific fact about $L(t)$ and $A(t)$ that will be needed is that $p_{2\nu}$ and $a_{j,2\nu-j-1}$ ($j = 0, 1, \cdots, \nu - 1$) are all equal to $\pm\omega_{\nu\nu}$. It follows then that if y is a solution there is an integer n such that

$$\int_a^b \langle x(t), L(t)\mathbf{y}^{(2\nu+1)}(t)\rangle dt = 0$$

for all $x \in O^{(n)}(a) \cap O^{(n)}(b)$. A simple argument shows that this implies that $L(t)\mathbf{y}^{(2\nu+1)}(t) \equiv 0$. [In the case $r = \omega$ use the fact that if this holds for all analytic x then it holds for all C^∞ x.] On the other hand $L(t)\mathbf{y}^{(2\nu+1)}(t) \equiv 0$ is a sufficient condition for y to be a solution, except for considerations of differentiability examined below. Now since $\omega_{\nu\nu}$ is non-degenerate, L is a differential operator of degree 2ν, hence $Ly \equiv 0$ has a unique solution for any prescribed value of $\mathbf{y}^{(2\nu)}(t)$. Now (a)–(d) follow from (1.2), and the fact that y is a solution if and only if $Ly \equiv 0$.

Differentiability. Since derivatives of order 2ν occur, the assumption that $2\nu \leqq r$ is necessary to carry out integration by parts. The coefficients of L involve ν^{th} derivatives of the ω's, hence are of class $C^{r-2\nu}$, which is just enough to make the solutions C^r, i.e. to guarantee that the solutions are in $V[a, b]$.

Note that L depends on the trivialization of $V[a, b]$, but the notion of solution does not.

1.3. *The hermitian form ψ on S.* For $y \in S$ the number $\psi[y] = \operatorname{Im}\langle \mathbf{y}^{(\nu)}(t), A(t)\mathbf{y}^{(2\nu)}(t)\rangle$ is independent of the choice of t and even of the choice of trivialization. To see this, observe that, for any $y \in S$, the map

$$x \to \int_{t_0}^{t_1} \Omega(t)[x, y]$$

of $V[a, b] \to \mathbf{C}$ depends, by (c), only on $\mathbf{x}^{(\nu)}(t_0)$ and $\mathbf{x}^{(\nu)}(t_1)$; even more, it depends on them separately. In this way property (c) of solutions can be used to show that there are unique maps $A(t)$: $S \to \{V[a, b]/O^{(\nu)}(t)\}^* = E^{\nu *}$ such that

$$\int_{t_0}^{t_1} \Omega(t)[x, y] = \langle \mathbf{x}^{(\nu)}(t), A(t)y\rangle \Big|_{t_0}^{t_1} \qquad (1.3)$$

for all $x \in V[a, b]$, $y \in S$. This is the invariant description of the A of (1.2). Now taking $x = y$ and using the fact that $\Omega(t)[y]$ is real, it follows that $\psi[y] = \operatorname{Im}\langle \mathrm{y}(t), A(t)y\rangle$ is independent of t. ψ is immediately seen to be a hermitian form on S, and we now claim that it is non-degenerate. To see this, define $A^\sharp(t)\colon S \to E^\nu \oplus E^{\nu *}$ by $A^\sharp(t)[y] = (\mathrm{y}(t), A(t)y)$. From $a_{j,2\nu-j-1} = \pm\omega_{\nu\nu}$ it is seen that the matrix of $A^\sharp(t)$ is non-singular, hence that $A^\sharp(t)$ is 1–1 onto. Now the non-degeneracy of ψ follows from the fact that for any finite dimensional complex vector space F the hermitian form $(v, w) \to \operatorname{Im}\langle v, w\rangle$ on $F \oplus F^*$ is non-degenerate.

The maps

$$A(t)\colon S \to E^{\nu *} \tag{1.4}$$
$$A^\sharp(t)\colon S \to E^\nu \oplus E^{\nu *} \qquad \text{1–1 onto}$$

together with the defining property (1.3) of A will be used frequently. Note that by (1.2) they are of class $C^{r-\nu+1}$ in t.

DEFINITION. By the *solution space* of Ω we mean the pair (S, ψ) consisting of the solutions S and the non-degenerate hermitian form ψ defined on them by $\psi[y] = \operatorname{Im}\langle \mathrm{y}(t), A(t)y\rangle$.

1.4. *d-Equivalence*. Let Ω be a derivative dependent hermitian form of order $\leqq \nu$ on $V[a, b]$. Define $d\Omega$ by $d\Omega(t)[x] = (d/ds)|_t\, \Omega(s)[x]$. Then (1.1) shows that $d\Omega$ is a derivative dependent hermitian form of order $\leqq \nu + 1$ (provided $2(\nu + 1) \leqq r$). Note that the coefficients are $C^{r-\nu-1}$ as required. Two derivative dependent hermitian forms will be called *d-equivalent* if their difference is d of something; what class this "something" should lie in is simplified by the fact, easily verified by formal integration by parts, that if Ω is non-degenerate of order ν, and Λ, β are derivative dependent hermitian forms of any order with Λ non-degenerate, then $\Omega - \Lambda = d\beta$ implies that order $\beta \leqq \nu - 1$, order Λ = order Ω, and Λ has the same leading coefficient as Ω. Thus the notions of *d-equivalence class of non-degenerate derivative dependent hermitian forms of order* ν and *d-equivalence class of Sturm forms of order* ν are well-defined, the difference of two members of the same class being d of a (not necessarily non-degenerate) form of lower order.

Now suppose that $\Omega - \Lambda = d\beta$ where Ω, Λ are non-degenerate of order ν. Then from the definition of solution, the fact that order $\beta < \nu$ and $\int_a^b d\beta(t)[x] = \beta(b)[x] - \beta(a)[x]$, it follows that Ω and Λ have the same solutions S. Moreover, by the defining property (1.3) we have $A_\Omega = A_\Lambda + \beta$ so, since β is real, the corresponding ψ's are equal. In short, Ω and Λ have the same solution space (S, ψ). Not only is the converse true, but we know exactly which pairs (S, ψ) arise:

THEOREM 1.2. *Two non-degenerate derivative dependent hermitian forms have the same solution space if and only if they are d-equivalent. Moreover,* (S, ψ) *is the solution space of some* Ω *if and only if:*

(a) $S \subset V[a, b]$ *is a subspace of dimension* $2\nu \cdot \dim E$ *where* $2\nu \leqq r$.

(b) $S \cap O^{(2\nu)}(t) = \{0\}$ *for all* $t \in [a, b]$.

(c) ψ *is a non-degenerate hermitian form on* S.

(d) *The restriction of* ψ *to* $S \cap O^{(\nu)}(t)$ *is* $\equiv 0$ *for all* $t \in [a, b]$.

1.5. PROOF OF THEOREM 1.2. We have already shown that "d-equivalent" implies "same solution space", and "(S, ψ) the solution space of some Ω" implies (a)–(d). The proof of the first statement is completed by:

LEMMA 1.3. *If* Ω_1 *and* Ω_2 *have the same solution space, then they are d-equivalent.*

PROOF. Let $A_j(t)\colon S \to E^{\nu *}$ for $j = 1, 2$ be the maps corresponding to Ω_1, Ω_2. Then if we have $\Omega_2 - \Omega_1 = d\beta$ the uniqueness of A in (1.3) gives

$$\beta(t)[x] = \beta(t)[y] = \langle \mathbf{y}^{(\nu)}(t), [A_2(t) - A_1(t)]y \rangle \tag{1.5}$$

for all $y \in S$ such that $x - y \in O^{(\nu)}(t)$. Thus we want to show that (1.5) can be used to *define* β. Since the solution spaces are assumed equal, the RHS is real. Moreover it is $C^{r-\nu+1}$ so we need only show that the β it defines (ostensibly of order $\leqq (2\nu - 1)$ since we must take $x - y \in O^{(2\nu)}(t)$ to define β) satisfies $\Omega_2 - \Omega_1 = d\beta$ (from which it follows that β is of order $\leqq \nu - 1$). Now $\mathbf{y}^{(\nu+1)}(t)$ for $y \in S$ gives a complete set of representatives of $V[a, b]/O^{(\nu+1)}(t)$ so it suffices to show that $\Omega_2(t)[y] - \Omega_1(t)[y] = d\beta(t)[y]$ for $y \in S$, and this is proved by using (1.3) to evaluate $\int_a^t \Omega_i(t)[y]$ and differentiating.

The following proposition proves that (S, ψ) arises from an Ω locally, as will be seen following its proof. It proves in addition the most difficult part of Theorem 2.1 below.

PROPOSITION 1.4. *Let* (S, ψ) *satisfy* (a)–(d) *of Theorem* 1.2, *and let* $\Pi \subset S$ *be given satisfying*:

(i) Π *is a subspace of dimension* $(1/2) \dim S = \nu \cdot \dim E$.

(ii) $\Pi \cap O^{(\nu)}(t) = \{0\}$ *for all* $t \in [a, b]$.

(iii) $\psi \mid \Pi \equiv 0$ ($\mid$ *denotes restriction*).

Then there is a unique non-degenerate derivative dependent hermitian form Ω *of order* ν *such that* (S, ψ) *is the solution space of* Ω, *and such that* $\Omega(t)[v, \pi] = 0$ *for all* $t \in [a, b]$, $v \in V[a, b]$, $\pi \in \Pi$.

REMARK. In terms of the notation of § 2, conditions (i)–(iii) can be restated: $\Pi \in \mathrm{U}(S, \psi)$ satisfies $\Phi(t) \cap \Pi = 0$ for all $t \in [a, b]$.

PROOF. We first prove uniqueness, finding in the process the formula for Ω in order to prove existence. Assume then that Ω satisfies the conclusion of the proposition. Then $\int_{t_0}^{t_1}\Omega(t)[v, \pi] = 0$ for all $v \in V[a, b]$, $\pi \in \Pi$, hence from (1.3) we obtain $A(t) \mid \Pi \equiv 0$. This turns out to be enough information to determine $A(t)$ in terms of Π and S. To see this, note first that for $x \in S$, we have $\operatorname{Im}\langle \mathbf{x}, Ax\rangle = \psi[x]$ while $\operatorname{Re}\langle \mathbf{x}(t), A(t)x\rangle = \operatorname{Re}\langle \mathbf{y}(t), A(t)z\rangle$ where $x = y + z$ is the unique decomposition of x with $y \in \Pi$ and $z \in O^{(\nu)}(t) \cap S$. Now let j_t: $S \rightarrow S$ be the unique automorphism of S which is the identity map on Π and multiplication by $-i$ on $O^{(\nu)}(t) \cap S$. Then

$$\begin{aligned}\psi[j_t x] &= \psi[y - iz] = \operatorname{Im}\langle \mathbf{y}(t) - i\mathbf{z}(t), A(t)(y - iz)\rangle \\ &= \operatorname{Im}\langle \mathbf{y}(t), -iA(t)z\rangle = \operatorname{Re}\langle \mathbf{y}(t), A(t)z\rangle \\ &= \operatorname{Re}\langle \mathbf{x}(t), A(t)x\rangle .\end{aligned}$$

Thus we have

$$\langle \mathbf{x}(t), A(t)x\rangle = \psi[j_t x] + i\psi[x] \tag{1.6}$$

for all $x \in S$. From this we have $\int_a^t \Omega(t)[x]$ for $x \in S$ in terms of S, ψ, and Π, and differentiating we have Ω in terms of S, ψ, and Π, which proves uniqueness.

We want to show then that given S, ψ, and Π, the formula (1.6) leads to an Ω. Writing both sides as sesqui-linear forms (in (1.6) they are *complex-valued* hermitian forms) we obtain

$$\langle \mathbf{x}(t), A(t)y\rangle = \psi[j_t x, j_t y] + i\psi[x, y] \tag{1.7}$$

for all $x, y \in S$ from which Ω is given by:

$$\Omega(t)[v_1, v_2] = \frac{d}{ds}\bigg|_t \psi[j_s x_1, j_s x_2] \tag{1.8}$$

where $v_i \in V[a, b]$ are arbitrary, and $x_i \in S$ are chosen such that $x_i - v_i \in O^{(\nu+1)}(t)$. We must show that (1.8) defines Ω

(1) independent of choices of x_i;
(2) non-degenerate;
(3) with solution space (S, ψ);
(4) satisfying the conclusion of the proposition;
(5) $C^{r-\nu}$.

The proof of these facts lies mainly in showing that (1.7) defines an A with the desired properties, namely that (1.7) *defines* $A(t)$: $S \rightarrow E^{\nu *}$ *onto with kernel* Π. To see this observe that for $x \in O^{(\nu)}(t) \cap S$, we have

$$\psi[j_t x, j_t y] + i\psi[x, y] = \psi[-ix, j_t y - y] .$$

For $y \in \Pi$, this is zero by the definition of j_t. For $y \in O^{(\nu)}(t) \cap S$, it is zero by assumption (d) of the theorem. Hence it is zero for all $y \in S$, and it follows that (1.7) depends only on the class of x mod $O^{(\nu)}(t)$; i.e., it indeed defines $A(t)\colon S \to E^{\nu *}$. If $A(t)$ is not onto, then its kernel is of dimension greater than $(1/2)\dim S$; taking the imaginary part, this gives a subspace of S of dimension greater than $(1/2)\dim S$ on which $\psi \equiv 0$, which is impossible if ψ is non-degenerate. Therefore $A(t)$ is onto by assumption (c). Finally, if $y \in \Pi$ we have

$$\psi[j_t x, j_t y] + i\psi[x, y] = \psi[j_t x + ix, y]$$

which is zero for $x \in O^{(\nu)}(t) \cap S$ by the definition of j_t, and is zero for $x \in \Pi$ by the assumption that $\psi \mid \Pi \equiv 0$. Hence Π is the kernel of $A(t)$.

(1) To show that (1.8) is independent of the choice of the x_i, it suffices by symmetry to show that it is independent of the choice of x_1. But it is $(d/ds)\,|_t \langle \mathbf{x}_1(s), A(s)x_2\rangle$, which, since $\mathbf{x}_1(s)$ depends on $(\nu - 1)$ derivatives, depends on only ν derivatives of x_1 at t.

(2) Suppose that $v \in O^{(\nu)}(t)$ is such that $\Omega(t)[v, w] = 0$ for all $w \in O^{(\nu)}(t)$. We want to conclude that $v \in O^{(\nu+1)}(t)$. Now taking $w \in O^{(\nu)}(t) \cap S$, we have

$$\begin{aligned} 0 = \Omega(t)[v, w] &= (d/ds)\,|_t \langle \mathbf{v}(s), A(s)w\rangle \\ &= \langle (d/ds)\,|_t\, \mathbf{v}(s), A(t)w\rangle\,. \end{aligned}$$

Since $A(t) \mid O^{(\nu)}(t) \cap S$ is onto $E^{\nu *}$, this means $(d/ds)\,|_t\, \mathbf{v}(s) = 0$; i.e., $v \in O^{(\nu+1)}(t)$.

(3) For $v \in V[a, b]$ and $y \in S$, we have $\Omega(t)[v, y] = (d/dt)\langle \mathbf{v}(t), A(t)y\rangle$ from which it follows that the elements of S are solutions. By non-degeneracy this must be all the solutions, and moreover for $y \in S$ we have $\operatorname{Im}\langle \mathbf{y}(t), A(t)y\rangle = \psi[y]$.

(4) For $y \in \Pi$, we have $A(t)y \equiv 0$.

(5) Since $t \to O^{(\nu)}(t) \cap S$ is $C^{r-\nu+1}$ (in the Grassmann manifold of subspaces of S of half the dimension) so is j_t (in the group of automorphisms of S), therefore so is A, from which Ω is $C^{r-\nu}$.

Now Ω's satisfying the condition of the proposition exist locally. To see this let $\mathrm{U}(S, \psi)$ be the manifold of all subspaces Π of S satisfying (i) and (iii) of the proposition. By (d) of the theorem $\mathrm{U}(S, \psi)$ is non-empty. The geometry of $\mathrm{U}(S, \psi)$ is studied extensively in § 4, and a trivial consequence of this study is that for any $t_0 \in [a, b]$ there is a $\Pi \in \mathrm{U}(S, \psi)$ such that $\Pi \cap [O^{(\nu)}(t_0) \cap S] = \{0\}$. Then there is a neighborhood of t_0 on which $\Pi \cap O^{(\nu)}(t) = \{0\}$ and on this neighborhood the proposition applies to give an Ω (on the neighborhood) with solution space (S, ψ). In this way we cover $[a, b]$ with a finite number of Ω's, which by Lemma (1.3) are d-equivalent on overlaps. The following lemma can then be used to piece these together to give a single Ω on $[a, b]$ with solution space (S, ψ), and hence to prove the theorem.

LEMMA 1.5. *Given real numbers $a < b < c < d$, and given Ω_1, Ω_2 on $V[a, c]$, $V[b, d]$ respectively, whose restrictions to $V[b, c]$ are d-equivalent, then there is an Ω on $V[a, d]$ whose restrictions to $V[a, c]$, $V[b, d]$ are d-equivalent to Ω_1, Ω_2.*

PROOF. If $r \neq \omega$, this is an easy consequence of the extendability of C^r functions defined on closed sets. In the general case, use Proposition (6.1) to conclude that L is unique on the interval and, from the same proposition, that there is then an Ω throughout the interval.

2. Index problems

2.1. *The index of the focal problem.*

DEFINITION. Let Q be a hermitian form on a complex vector space V. Consider the set of all finite dimensional subspaces of V on which Q is negative definite. If the elements of this set are bounded in dimension, we call the maximum dimension the *index* of Q, denoted $\mathrm{ix}(Q)$. Otherwise we set $\mathrm{ix}(Q) = \infty$. The *nullity* of Q, also a non-negative integer or ∞, is defined to be the dimension of $V^{\perp}(\mathrm{rel.}\ Q)$.

REMARKS. 1. If V is finite dimensional, then a basis $V \approx \mathbf{C}^n$ can be chosen in which $Q[(x_1, \cdots, x_n)] = \sum_{i=1}^{n} \varepsilon_i \,|\, x_i \,|^2$ where ε_i is ± 1 or 0. Then $\mathrm{ix}(Q)$ is the number of -1's, and $\mathrm{nul}(Q)$ is the number of 0's.

2. Q is non-degenerate if and only if $\mathrm{nul}(Q) = 0$.

3. It is easily shown that if $\mathrm{nul}(Q) = \mathrm{ix}(Q) = 0$, then Q is positive definite.

For the moment we will be interested in the case where $V = O^{(\nu)}(a) \cap O^{(\nu)}(b)$, and Q is $\int_a^b \Omega$ where Ω is a non-degenerate derivative dependent hermitian form of order ν. This hermitian form on this space will be denoted, for reasons explained in 2.2, by $(\Omega, \infty, [a, b])$. Note that $(\Omega, \infty, [a, b])$ depends only on the d-equivalence class of Ω, which in turn depends only on the solution space (S, ψ). The main result of this paragraph expresses $\mathrm{ix}(\Omega, \infty, [a, b])$ in terms of (S, ψ). The expression of $\mathrm{nul}(\Omega, \infty, [a, b])$ in terms of (S, ψ) is elementary—from (c) of Proposition 1.1 we have $\mathrm{nul}(\Omega, \infty, [a, b]) = \dim [S \cap O^{(\nu)}(a) \cap O^{(\nu)}(b)] = \dim [\Phi(a) \cap \Phi(b)]$ where $\Phi(t) \subset S$ is defined by $\Phi(t) = S \cap O^{(\nu)}(t)$. The expression of $\mathrm{ix}(\Omega, \infty, [a, b])$ is obtained from the following two theorems:

LOCAL INDEX THEOREM 2.1. *Let Ω be a non-degenerate derivative dependent hermitian form of order ν. Then the following are equivalent:*

(a) $\mathrm{ix}(\Omega, \infty, [a, b]) < \infty$.

(b) *Setting $\Phi(t) = S \cap O^{(\nu)}(t)$, the curve $t \to \Phi(t)$ in $\mathrm{U}(S, \psi)$ is a $\oplus$-curve.*

(c) *Ω is a Sturm form,* i.e., *its leading coefficient is positive definite.*

(d) $(\Omega, \infty, [a, b])$ *is locally positive definite, by which we mean that there is an* $\varepsilon > 0$ *such that* $(\Omega, \infty, [t_0, t_1])$ *is positive definite whenever* $[t_0, t_1] \subset [a, b]$ *is a subinterval of length* $< \varepsilon$.

REMARKS. 1. (b) is to be regarded as the "expression in terms of (S, ψ)" of this property.

2. In view of (a) only Sturm forms are of interest in index problems. The rest of the paper deals exclusively with Sturm forms.

MORSE INDEX THEOREM 2.2. *If* Ω *is a Sturm form, then*

$$\mathrm{ix}(\Omega, \infty, [a, b]) = \sum_{a<t<b} \mathrm{nul}(\Omega, \infty, [a, t])$$

or equivalently

$$= \sum_{a<t<b} \dim [\Phi(a) \cap \Phi(t)] .$$

PROOF OF 2.1. (b) $\Leftrightarrow$ (c). Let $t_0 \in [a, b]$ be given. Choose $\Pi \in \mathrm{U}(S, \psi)$ such that $\Pi \cap \Phi(t_0) = \{0\}$. The tangent vector to $t \to \Phi(t)$ is the form $(d/dt)|_{t_0} \alpha[\Pi, \Phi(t_0), \Phi(t)]$ which, for any $v \in \Phi(t_0)$, has the value $(d/dt)|_{t_0} \psi[j_t v]$ where j_t is multiplication by 1 on Π, and $-i$ on $\Phi(t)$. By (1.8) this is $\Omega_\Pi(t_0)[v]$, where Ω_Π is the form d-equivalent to Ω given by Proposition 1.4. Therefore, since $v^{(\nu)}(t_0) = 0$, the tangent vector evaluated on v is equal to the leading coefficient of Ω (since Ω, Ω_Π have the same leading coefficient) evaluated on $v^{(\nu)}(t_0)$. It follows that the tangent vector is positive semi-definite if and only if (c) is satisfied. Thus by (III) of Proposition 4.7, it suffices to show that $t \to \Phi(t)$ is discretely self-intersecting. This is done by a classical argument using the fact (Proposition 1.1 (b)) that the wronskian of S is everywhere non-zero, the underlying idea being that any solution which is zero at 2ν nearby (or coincident) points is identically zero.

(c) $\Rightarrow$ (d). We will show that *if* (c) *is satisfied, and if there is a* Π *as in Proposition* 1.4, *then* $(\Omega, \infty, [a, b])$ *is positive definite.* To deduce (c) $\Rightarrow$ (d) from this statement, let I be the collection of subintervals $[t_0, t_1] \subset [a, b]$ with the property that there is a $\Pi \in \mathrm{U}(S, \psi)$ for which $\Phi(t) \cap \Pi = \{0\}$ when $t \in [t_0, t_1]$. We want to show that there is an $\varepsilon > 0$ such that all subintervals of length $< \varepsilon$ are in I. This is true because:

(i) subintervals of elements of I are in I, and

(ii) every point of $[a, b]$ is interior to an element of I.

Thus let $\Pi \in \mathrm{U}(S, \psi)$ be given with $\Pi \cap \Phi(t) = \{0\}$ for $t \in [a, b]$. Since $(\Omega, \infty, [a, b])$ depends only on the d-equivalence class we can assume from Proposition 1.4 that $\Omega(t)[v, \pi] = 0$ for $\pi \in \Pi$, $v \in V[a, b]$. Then $\Omega(t)[v] = \Omega(t)[v - \pi]$. Now let L_Π be the unique ν^{th} order differential operator, with leading coefficient the identity whose solution space is Π. Explicitly $L_\Pi(t)[v]$ is the ν^{th} derivative of $v - \pi$ at t where π is the unique element of Π for which $v - \pi \in O^{(\nu)}(t)$. Then $\Omega(t)[v]$ is equal to the leading coefficient of Ω evaluated on $L_\Pi v$. Hence by (c) we have $\Omega(t)[v] \geqq 0$ with equality if and only if $L_\Pi(t)[v] = 0$. Therefore $\int_a^b \Omega(t)[v] \geqq 0$ with equality if and only if $L_\Pi v \equiv 0$, i.e., $v \in \Pi$. Thus $v \in$ domain $(\Omega, \infty, [a, b])$ and $\int_a^b \Omega(t)[v] = 0 \Rightarrow v = 0$ and (d) is proved.

(d) $\Rightarrow$ (a). Let $\tau = (t_0, t_1, \cdots, t_n, t_{n+1})$ be a subdivision $a = t_0 < t_1 < \cdots < t_n < t_{n+1} = b$ of $[a, b]$ with the property that $\Phi(t_i) \cap \Phi(t_{i+1}) = \{0\}$ for $i = 0, 1, \cdots, n$. For such a τ, let $B(\tau)$ be the space of broken solutions of Ω relative to the subdivision τ, by which we mean

$$B(\tau) = \{(x_0, x_1, \cdots, x_n) \in S^{n+1} : x_0 \in O^{(\nu)}(t_0);\ x_i - x_{i-1} \in O^{(\nu)}(t_i) \text{ for } i = 1, \cdots, n;\ x_n \in O^{(\nu)}(t_{n+1})\}\ ,$$

and let $Q(\tau)$ be the hermitian form $\int_a^b \Omega$ on $B(\tau)$, by which we mean

$$Q(\tau)[(x_0, \cdots, x_n)] = \sum_{i=0}^n \int_{t_i}^{t_{i+1}} \Omega(t)[x_i]\ .$$

From $\Phi(t_i) \cap \Phi(t_{i+1}) = \{0\}$, it follows that ev: $B(\tau) \to E^{n\nu}$ defined by $\mathrm{ev}[(x_0, x_1, \cdots, x_n)] = (\mathbf{x}_1(t_1), \mathbf{x}_2(t_2), \cdots, \mathbf{x}_n(t_n))$ is 1–1 onto. Hence we can define p: $O^{(\nu)}(a) \cap O^{(\nu)}(b) \to B(\tau)$ by $p(v) = \mathrm{ev}^{-1}[\mathbf{v}(t_1), \mathbf{v}(t_2), \cdots, \mathbf{v}(t_n)]$. In short $p(v)$ is the broken solution which agrees with v at the points of τ.

Now choose τ such that $t_{i+1} - t_i < \varepsilon$ where ε is as in (d). Then it follows that $\Phi(t_i) \cap \Phi(t_{i+1}) = \{0\}$, and that $\int_a^b \Omega(t)[v] \geqq Q(\tau)[p(v)]$. Therefore if $W \subset O^{(\nu)}(a) \cap O^{(\nu)}(b)$ is a subspace on which $\int_a^b \Omega$ is negative definite, then $p \mid W$ has no kernel, and $Q(\tau)$ is negative definite on $p[W] \subset B(\tau)$. Hence $\mathrm{ix}(\Omega, \infty, [a, b]) \leqq \mathrm{ix}[Q(\tau)] < \infty$ and (a) is proved.

(a) $\Rightarrow$ (c). It will suffice to show that if (c) does not hold, then there is an infinite dimensional subspace of $O^{(\nu)}(a) \cap O^{(\nu)}(b)$ on which $\int_a^b \Omega$ is negative definite. Now if $r \neq \omega$, we can argue as follows: If (c) does not hold then we can choose $F \subset E$ and $[t_0, t_1] \subset [a, b]$ such that the leading coefficient of Ω is negative definite on F for $t \in [t_0, t_1]$. Let $V_F[t_0, t_1] \subset V[t_0, t_1]$ be the subspace consisting of curves whose image is in F (i.e., sections of the subbundle $F \times [t_0, t_1] \subset E \times [t_0, t_1]$). Then by choosing t_0, t_1 closer together if necessary we have by (c) $\Rightarrow$ (d) that $\int_{t_0}^{t_1} (-\Omega)$ is positive definite on $V_F[t_0, t_1] \cap O^{(\nu)}(t_0) \cap O^{(\nu)}(t_1) \subset V_F[t_0, t_1] \cap O^{(r+1)}(t_0) \cap O^{(r+1)}(t_1)$. Extending elements of the latter space to elements of $V[a, b]$ by making them identically zero on $[a, t_0)$ and $(t_1, b]$, we have the desired result. In the general case it is probably easiest to decompose $E \times [a, b]$ into the sum of two subbundles such that Ω is a Sturm form on one, and $-\Omega$ is a Sturm form on the other. Then except for a finite dimensional subspace, $\int_a^b \Omega$ is negative definite on the space of sections on the second bundle.

This completes the proof of Theorem 2.1.

The full proof of Theorem 2.2 is postponed to § 3 where it is proved as a part of Theorem 3.1. An important step in the proof of Theorem 3.1 is the following partial statement of Theorem 2.2.

PROPOSITION 2.4. *If* Ω *is a Sturm form and if* $\mathrm{nul}(\Omega, \infty, [a, b]) = 0$, *then we have*

$$\mathrm{ix}(\Omega, \infty, [a, b]) \leqq \sum_{a<t<b} \mathrm{nul}(\Omega, \infty, [a, t]) .$$

PROOF. This is an extension of the proof of (d) $\Rightarrow$ (a) above. In terms of the notation of that proof, it will suffice to show:

PROPOSITION 2.4′. *If* Ω *is a Sturm form, and if* $\Phi(a) \cap \Phi(b) = \{0\}$, *then there is a subdivision* τ *of* $[a, b]$ *for which*

$$\mathrm{ix}[Q(\tau)] = \sum_{a<t<b} \dim [\Phi(a) \cap \Phi(t)] .$$

PROOF. Since $\Phi(a) \cap \Phi(b) = \{0\}$, we can choose $\tau = (t_0, \cdots, t_{n+1})$ such that $t_i - t_{i-1} < \varepsilon$, and $\Phi(t_0) \cap \Phi(t_i) = \{0\}$ for $i = 1, 2, \cdots, n+1$, where ε is as in (d) of Theorem 2.1. Note that it follows that $\Phi(t) \cap \Phi(t_i) = \{0\}$ for $t \in [t_{i-1}, t_{i+1}]$, $t \neq t_i$. We will show that the proposition holds for this τ. To this end, let $\tau_1 = (t_0, \cdots, t_{n-1}, t_n)$ and $\tau_2 = (t_0, t_n, t_{n+1})$. Then there are natural inclusions of $B(\tau_1)$ and $B(\tau_2)$ in $B(\tau)$ (the assumption that $\Phi(t_0) \cap \Phi(t_n) = \{0\}$ is needed for $B(\tau_2)$ to be defined) whose images we identify with $B(\tau_1)$ and $B(\tau_2)$. Then $B(\tau_1) \cap B(\tau_2) = \{0\}$, and by dimensionality we have $B(\tau) = B(\tau_1) \oplus B(\tau_2)$. Now the elements of $B(\tau_2)$ are unbroken solutions on $[t_0, t_n]$, and the elements of $B(\tau_1)$ are identically zero on $[t_n, t_{n+1}]$; so from (1.3) we have easily that $B(\tau_1) \perp B(\tau_2)$ (rel. $Q(\tau)$). Hence $B(\tau_1) \oplus B(\tau_2)$ splits $Q(\tau)$ into $Q(\tau_1) \oplus Q(\tau_2)$ and both steps of an inductive argument will be proved if it is shown that $\mathrm{ix}[Q(\tau_2)] = \sum_{t_n<t<t_{n+1}} \dim [\Phi(t) \cap \Phi(a)]$. For this, consider the index of $\alpha[\Phi(t_0), \Phi(t), \Phi(t_{n+1})]$. From (I) of Proposition 4.7, it is zero for $t = t_{n+1} - \delta$ for small δ, hence from Proposition 4.5, we have $\mathrm{ix}\alpha[\Phi(t_0), \Phi(t_n), \Phi(t_{n+1})] = \sum_{t_n<t<t_{n+1}} \dim [\Phi(a) \cap \Phi(t)]$. The proof of the proposition will thus be complete once we prove:

LEMMA 2.5. *If* τ *consists of three values* (t_0, t_1, t_2), *and if* $\Phi(t_0)$, $\Phi(t_1)$, $\Phi(t_2)$ *are mutually disjoint, then the hermitian forms* $Q(\tau)$ *on* $B(\tau)$ *and* $\alpha[\Phi(t_0), \Phi(t_1), \Phi(t_2)]$ *on* $\Phi(t_1)$ *are equivalent.*

REMARK. This shows the connection between (b) and (d) of Theorem 2.1.

PROOF. From $\Phi(t_0) \cap \Phi(t_2) = \{0\}$ it follows that $(x_0, x_1) \rightarrow x_0 - x_1$ carries

$B(\tau)$ one-one onto $\Phi(t_1)$. We will show that this induces an equivalence of the two forms. We have

$$Q(\tau)[(x_0, x_1)] = \langle \mathbf{x}_0(t), A(t)x_0\rangle \Big|_{t_0}^{t_1} + \langle \mathbf{x}_1(t), A(t)x_1\rangle \Big|_{t_1}^{t_2} .$$

From $x_0 \in \Phi(t_0)$ and $x_1 \in \Phi(t_2)$ the terms at t_0 and t_2 are zero while the terms at t_1 are both real. Thus

$$Q(\tau)[(x_0, x_1)] = \mathrm{Re}\{\langle \mathbf{x}_0(t_1), A(t_1)x_0\rangle - \langle \mathbf{x}_1(t_1), A(t_1)x_1\rangle\} .$$

Using the fact that $x_0(t_1) = x_1(t_1)$, this can be written

$$\begin{aligned} &= \mathrm{Im}\{\langle i\mathbf{x}_1(t_1), A(t_1)x_0\rangle + \langle \mathbf{x}_0(t_1), A(t_1)ix_1\rangle\} \\ &= \mathrm{Im}\{\langle \mathbf{x}_0(t_1) + i\mathbf{x}_1(t_1), A(t_1)(x_0 + ix_1)\rangle\} \\ &= \psi[x_0 + ix_1] = \psi[j(x_0 - x_1)] \end{aligned}$$

where $j\colon S \to S$ is multiplication by 1 on $\Phi(t_0)$ and by $-i$ on $\Phi(t_2)$. Since this is $\alpha[\Phi(t_0), \Phi(t_1), \Phi(t_2)]$ evaluated on $x_0 - x_1$, the lemma is proved.

2.2. *The general index problem.* The generalized Sturm theorem of § 3 deals with hermitian forms of the following type:

DEFINITION. Let Ω be a Sturm form of order ν on $V[a, b]$, and let $\beta \in \mathrm{Herm}(E^\nu \oplus E^\nu)$. We denote by $(\Omega, \beta, [a, b])$ the hermitian form defined on

$$\{v \in V[a, b] : (\mathbf{v}^{(\nu)}(a), \mathbf{v}^{(\nu)}(b)) \in \text{domain } \beta\}$$

by

$$v \to \int_a^b \Omega(t)[v] - \beta[(\mathbf{v}^{(\nu)}(a), \mathbf{v}^{(\nu)}(b))] .$$

$(\Omega, \beta, [a, b])$ is thus a particular kind of hermitian pair on $V[a, b]$.

EXAMPLES. 1. If $\beta = \infty \in \mathrm{Herm}(E^\nu \oplus E^\nu)$, then $(\Omega, \infty, [a, b])$ has its former meaning.

2. If $\beta = \gamma_1 \oplus \gamma_2$ where $\gamma_i \in \mathrm{Herm}(E^\nu)$ (i.e., domain β = domain $\gamma_1 \oplus$ domain γ_2 and $\beta[(x_1, x_2)] = \gamma_1[x_1] + \gamma_2[x_2]$), then β is said to be *separated*. In the calculus of variations such β's arise from the problem of an extremal subject to variation through paths whose end points lie on two given submanifolds (see § 8).

3. Domain β is the diagonal of $E^\nu \oplus E^\nu$ and $\beta[(x, x)] = 0$. This is the *periodic* boundary condition. In the calculus of variations it arises from the problem of a closed extremal subject to variation through closed paths.

4. $\beta = 0$. Then $(\Omega, \beta, [a, b])$ is simply $\int_a^b \Omega$ on $V[a, b]$.

We note in passing that $(\Omega, \beta, [a, b])$ depends on more than the d-equivalence class of Ω; but that if Ω_1 and Ω_2 are d-equivalent, then there is an automorphism $\beta_1 \to \beta_2$ of $\mathrm{Herm}(E^\nu \oplus E^\nu)$ such that $(\Omega_1, \beta_1, [a, b]) = (\Omega_2, \beta_2, [a, b])$.

We want to compute $\mathrm{ix}(\Omega, \beta, [a, b])$. The advantage of the present formulation is that it leads immediately to the following reduction of this problem:

PROPOSITION 2.6. *Let* Ω *on* $V[a, b]$ *be such that* $\mathrm{nul}(\Omega, \beta, [a, b]) = 0$; i.e., *such that* $x \to (\mathrm{x}(a), \mathrm{x}(b))$ *maps* $S \to E^\nu \oplus E^\nu$ *non-singularly. Define a hermitian form* α *on* $E^\nu \oplus E^\nu$ *by* $\alpha(\mathrm{x}(a), \mathrm{x}(b)) = \int_a^b \Omega(t)[x]$ *for all* $x \in S$. *Then for any* $\beta \in \mathrm{Herm}(E^\nu \oplus E^\nu)$ *we have*

$$\mathrm{nul}(\Omega, \beta, [a, b]) = \mathrm{nul}(\Omega, \infty, [a, b]) + \mathrm{nul}(\alpha - \beta) = \mathrm{nul}(\alpha - \beta)\,,$$
$$\mathrm{ix}(\Omega, \beta, [a, b]) = \mathrm{ix}(\Omega, \infty, [a, b]) + \mathrm{ix}(\alpha - \beta)$$

where we take domain $(\alpha - \beta) =$ *domain* (β).

PROOF. This is simply the observation that the decomposition $V[a, b] = [O^{(\nu)}(a) \cap O^{(\nu)}(b)] \oplus S$ splits $\int_a^b \Omega$, hence $(\Omega, \beta, [a, b])$ is split into $(\Omega, \infty, [a, b])$ and $x \to \alpha[x] - \beta[\mathrm{x}(a), \mathrm{x}(b)]$ for $x \in S \cap$ domain $(\Omega, \beta, [a, b])$.

COROLLARY. *In this case* $\mathrm{ix}(\Omega, \beta, [a, b]) < \infty$. [It is shown in Theorem 3.1 that this is true in all cases].

Taken together with Theorem 2.2 this proposition gives the rule for computing $\mathrm{ix}(\Omega, \beta, [a, b])$ given just S (ψ is irrelevant if we know Ω is a Sturm form) and the value of $\int_a^b \Omega$ on S. The case $\mathrm{nul}(\Omega, \infty, [a, b]) \neq 0$ is excluded (see § 8 for the simple extension of the algorithm to this case) and is not of interest here, but it is important in § 3 to have the following extension of the definition of α to all cases:

DEFINITION 2.7. Let $F = E^\nu \oplus E^{\nu*} \oplus E^\nu \oplus E^{\nu*}$, and let Ψ be the hermitian form on F defined by $\Psi(v_1, w_1, v_2, w_2) = -\mathrm{Im}\langle v_1, w_1\rangle + \mathrm{Im}\langle v_2, w_2\rangle$. Then $(v_1, w_1, v_2, w_2) \to (v_1, v_2, -w_1, w_2)$ induces an identification $\mathrm{U}(F, \Psi) \approx \mathrm{Herm}(E^\nu \oplus E^\nu)$. Now if Ω is a Sturm form on $V[a, b]$, then the image of the map $A^\#(a) \oplus A^\#(b)$: $S \to F$ is an element of $\mathrm{U}(F, \Psi)$. We denote the corresponding element of $\mathrm{Herm}(E^\nu \oplus E^\nu)$ by $\alpha(\Omega, [a, b])$.

It is easily checked using Proposition 4.1 that $\alpha(\Omega, [a, b])$ can also be described by:

$$\text{domain } \alpha(\Omega, [a, b]) = \{(\mathrm{x}(a), \mathrm{x}(b)) \in E^\nu \oplus E^\nu : x \in S\}$$
$$\alpha(\Omega, [a, b])[(\mathrm{x}(a), \mathrm{x}(b))] = \int_a^b \Omega(t)[x] \qquad \text{for all } x \in S$$

so Definition 2.7 is an extension of that of Proposition 2.6. It is also easily checked that:

PROPOSITION 2.8. $\mathrm{nul}(\Omega, \beta, [a, b]) = \dim[\alpha(\Omega, [a, b]) \cap \beta]$.

3. The generalized Sturm theorem

THEOREM 3.1. *Let Ω be a Sturm form of order ν on $V[a, b]$, and let $\beta \in \mathrm{Herm}(E^\nu \oplus E^\nu)$. Then $\mathrm{ix}(\Omega, \beta, [a, b]) < \infty$ and:*

If Λ is any Sturm form on $V[a, b]$ for which $\int_a^b \Lambda$ is positive definite then

(A) $$\mathrm{ix}(\Omega, \beta, [a, b]) = \sum_{\lambda>0} \mathrm{nul}(\Omega + \lambda\Lambda, \beta, [a, b]) .$$

In the case where β is of the form $\beta = \gamma \oplus \infty$ where $\gamma \in \mathrm{Herm}(E^\nu)$, then (A) *is also equal to*

(B) $$\sum_{a<t<b} \mathrm{nul}(\Omega, \gamma \oplus \infty, [a, t]) .$$

Finally, setting $\alpha(\lambda) = \alpha(\Omega + \lambda\Lambda, [a, b])$ for $\lambda > 0$, (A) *is also equal to*

(C) $$[\alpha(\lambda) : \Gamma_\beta]_{\varepsilon \leqq \lambda \leqq \lambda^+}$$

whenever ε is sufficiently small and λ^+ sufficiently large.

REMARKS. 1. The normal choice of Λ (i.e., the choice which leads to familiar theorems) is a 0^{th} order form $\Lambda(t)[x] = \|x(t)\|^2$ where $\| \ \|$ is a hermitian norm on E.

2. Note that if $\lambda > 0$, then $\Omega + \lambda\Lambda$ is again a Sturm form, its order being max{order Λ, order Ω}. If order $\Lambda >$ order Ω then $(\Omega + \lambda\Lambda, \beta, [a, b])$ has not yet been defined. Observe, however, that the definition of $(\Omega, \beta, [a, b])$ in no way depends on the fact that $\nu =$ order Ω, so it is clear how to make the extension to the case $\nu \neq$ order Ω.

3. The Morse index theorem is the equality (B) = LHS(A) when $\beta = \gamma \oplus \infty$. Note that this includes Theorem 2.2.

4. The Sturm oscillation theorem is the equality (B) = RHS(A) when $\beta = \gamma \oplus \infty$. The other Sturm theorems are deduced from Theorem 3.1 in § 7.

5. The usefulness of (C) is that it puts (A) in a form which behaves well under perturbations. For example it easily yields:

COROLLARY. *Let $s \to \Omega_s$ be a family of Sturm forms, all of the same degree, depending smoothly on the real parameter s near 0. [Smoothly here means all terms of* (1.2) *are continuous in s.] Then if $\mathrm{nul}(\Omega_0, \beta, [a, b]) = 0$ we have $\mathrm{ix}(\Omega_s, \beta, [a, b]) = \mathrm{ix}(\Omega_0, \beta, [a, b])$ for all s sufficiently near 0.*

The following lemma will be proved as the last step of the proof of the theorem:

We will say that a Sturm form Λ on $V[a, b]$ has "property $\oplus$" if $\int_a^b \Lambda$ is positive definite and if $(\Lambda, \infty \oplus 0, [t, b])$ is positive definite (on its domain) for all $t \in [a, b)$.

LEMMA 3.2. *If Ω, Λ are Sturm forms on $V[a, b]$ and Λ has property $\oplus$, then $\Omega + \lambda\Lambda$ has property $\oplus$ for all sufficiently large λ.*

Note. From (B) of the theorem (with the orientation reversed as in (3.1) below) it follows easily that "$\int_a^b \Lambda$ positive definite" implies "Λ has property $\oplus$", so this concept is purely a stopgap. Notice then that the lemma is an immediate consequence of (A) [take $\beta = 0$].

PROOF OF THEOREM 3.1. The proof is given in steps (I)–(IX).

(I). *If order $\Lambda \leqq$ order Ω, and if Λ has property $\oplus$ then*

$$\sum_{\lambda>0} \mathrm{nul}(\Omega + \lambda\Lambda, \infty \oplus \gamma, [a, b]) = \sum_{a<t<b} \mathrm{nul}(\Omega, \infty \oplus \gamma, [t, b]) . \tag{3.1}$$

In other words, (B) = RHS(A) where we have changed the orientation of $[a, b]$ to make the signs more convenient. The method of proof is a direct generalization of the proof of the oscillation theorem in the introduction. For $\lambda \geqq 0$, let $(S_\lambda, \psi_\lambda)$ be the solution space of $\Omega + \lambda\Lambda$, and let $A^\#(\lambda, t)$: $S \to E^\nu \oplus E^{*\nu}$ be the map of (1.4). The proof will result from an examination of the map c: $[0, \infty) \times [a, b] \to \mathrm{Herm}(E^\nu)$ given by

$$c(\lambda, t) = A^\#(\lambda, b)A^\#(\lambda, t)^{-1}[\infty] .$$

For each fixed λ, this is the image under $A^\#(\lambda, b)$ of the curve $t \to \Phi(t)$, hence $t \to c(\lambda, t)$ is a $\oplus$-curve for each fixed λ by (b) of Theorem 2.1. Next we claim that, for $t \neq b$, the curve $\lambda \to c(\lambda, t)$ is also a $\oplus$-curve. This is proved by computing the tangent vector as follows:

The ψ defining $\mathrm{Herm}(E^\nu)$ corresponds to the hermitian sesqui-linear form $\psi[(u_1, v_1), (u_2, v_2)] = (1/2i)[\langle u_1, v_2\rangle - \overline{\langle u_2, v_1\rangle}]$. Now let $\lambda \to (u_\lambda, v_\lambda) \in c(\lambda, t)$ be a differentiable curve, say $(u_\lambda, v_\lambda) = A^\#(\lambda, b)[x_\lambda]$ where $x_\lambda \in S_\lambda \cap O^{(\nu)}(t)$. Then by the corollary of Proposition 4.6, the tangent vector at $\lambda = \xi$ evaluated on (u_ξ, v_ξ) is

$$\begin{aligned}
&\frac{d}{d\lambda}\Big|_\xi 2 \,\mathrm{Im}\, \psi[(u_\lambda, v_\lambda), (u_\xi, v_\xi)] \\
&= \frac{d}{d\lambda}\Big|_\xi \mathrm{Re}\, \{\langle \mathbf{x}_\xi(b), A(\lambda, b)x_\lambda\rangle - \langle \mathbf{x}_\lambda(b), A(\xi, b)x_\xi\rangle\} \\
&= \frac{d}{d\lambda}\Big|_\xi \mathrm{Re}\left\{\int_t^b (\Omega + \lambda\Lambda)[x_\xi, x_\lambda] - \int_t^b (\Omega + \xi\Lambda)[x_\lambda, x_\xi]\right\} \\
&= \frac{d}{d\lambda}\Big|_\xi \mathrm{Re}\left\{(\lambda - \xi)\int_t^b \Lambda[x_\xi, x_\lambda]\right\} = \int_t^b \Lambda[x_\xi]
\end{aligned}$$

which is positive by the hypothesis that Λ has property $\oplus$.

Thus (3.1) can be rewritten as

$$[c(\lambda, a) : \Gamma_\gamma]_{\lambda>0} = [c(0, t) : \Gamma_\gamma]_{a<t<b} \tag{3.2}$$

using the facts that the curves are $\oplus$-curves and $\dim [c(\lambda, t) \cap \gamma] =$

$\text{nul}(\Omega + \lambda\Lambda, \infty \oplus \gamma, [t, b])$. To prove (3.2) we show first that there are numbers t_0, λ_0 such that $c(\lambda, t) \cap \gamma = \{0\}$ if either $t \in [t_0, b)$ or $\lambda \in [\lambda_0, \infty)$. To find t_0, first use Theorem 2.1 to choose $t_1 \in [a, b)$ with the property that $(\Omega, \infty, [t, b])$ is positive definite for all $t \geqq t_1$. It follows that $c(\lambda, t) \cap \infty = \{0\}$ for all $t \in [t_1, b)$ and $\lambda \geqq 0$. Then choosing $t_0 \geqq t_1$ such that $c(0, t) \cap \gamma = \{0\}$ for $t \in [t_0, b)$, it follows, using (I) and (II) of Proposition 4.7 and the fact that $c(0, b) = \infty$, that $c(\lambda, t) \cap \gamma = \{0\}$ whenever $t \in [t_0, b)$. To find λ_0 choose a derivative dependent hermitian form δ of order $<$ order Ω with the property that $\gamma - \delta(b)$ is positive definite on domain γ [considering $\delta(b)$ as a hermitian form on E^ν]. Then choosing λ_0 such that $(\Omega + d\delta) + \lambda_0\Lambda$ has property $\oplus$ it follows that $(\Omega + \lambda\Lambda, \infty \oplus \gamma, [t, b])$ is positive definite for $\lambda \geqq \lambda_0$, from which $c(\lambda, t) \cap \gamma = \{0\}$ for $\lambda \geqq \lambda_0$.

Now consider $c(\lambda, t)$ as (λ, t) runs around the boundary of the rectangle $[0, \lambda_0] \times [a, t_0]$. The total intersection with Γ_γ is zero (the image curve is null-homotopic) and there is no intersection on two of the sides, from which (provided $c(0, a) \notin \Gamma_\gamma$) (3.2) follows. For the general case cut the corner $[0, \delta_1] \times [a, a + \delta_2]$ out of the rectangle. It is easy to do this in such a way that the only intersections with Γ_γ on the boundary of the small rectangle are at $(0,a)$ and along the open segment $\lambda = \delta_1$, $t \in (a, a + \delta_2)$. Then, following the boundary of the big rectangle minus the corner, we have LHS(3.2) $\geqq$ RHS(3.2). The opposite inequality is obtained similarly and (I) is proved.

(II). (C) *follows from* (A). By Proposition 2.8, it suffices to show that $\lambda \to \alpha(\lambda)$ is a $\oplus$-curve in $\text{Herm}(E^\nu \oplus E^\nu)$. This is done by computing the tangent vector exactly as in the proof above that $\lambda \to c(\lambda, t)$ is a $\oplus$-curve.

(III). $\geqq$ *holds in* (A). This is a standard argument. (Moreover, it is valid for hermitian forms in general and a setting such as that of Proposition 5.1 can be used.) Let $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_n$ be values for which $\text{nul}(\Omega + \lambda_i\Lambda, \beta, [a, b]) \neq 0$, and let $N_i \subset V[a, b]$ be the corresponding null-spaces. Then for $y_i \in N_i$, $y_j \in N_j$, we have

$$\int_a^b (\Omega + \lambda_i\Lambda)[y_i, y_j] - \beta[\hat{y}_i, \hat{y}_j] = 0 = \int_a^b (\Omega + \lambda_j\Lambda)[y_i, y_j] - \beta[\hat{y}_i, \hat{y}_j]$$

where we have abbreviated $(\mathbf{y}_i^{(\nu)}(a), \mathbf{y}_i^{(\nu)}(b))$ by $\hat{y}_i$. Thus

$$-\lambda_i\int_a^b \Lambda[y_i, y_j] = \int_a^b \Omega[y_i, y_j] - \beta[\hat{y}_i, \hat{y}_j] = \lambda_j\int_a^b \Lambda[y_i, y_j] .$$

And if $\lambda_i \neq \lambda_j$, all terms must be zero. Hence $(\Omega, \beta, [a, b])$ splits on $N_1 \oplus \cdots \oplus N_n$ (noting that the formula above also proves that $N_i \cap N_j = \{0\}$ for $i \neq j$). By the positivity of Λ, it is negative definite on each N_i separately, hence $\text{ix}(\Omega, \beta, [a, b]) \geqq \dim (N_1 \oplus \cdots \oplus N_n)$.

(IV). LHS(A) = RHS(A) = (B) *in the case where* $\beta = \infty$, $\gamma = \infty$, *order* $\Lambda \leqq$ *order* Ω, Λ *has property* $\oplus$, *and* $\mathrm{nul}(\Omega, \infty, [a, b]) = 0$. From (I) and (III) we have LHS(A) $\geqq$ RHS(A) = (B). But we have (B) $\geqq$ LHS(A), by Proposition 2.4.

(V). *The assumption "*Λ *has property* $\oplus$*" is superfluous.* Assume Ω has the property that $\int_a^b \Omega$ is positive definite on $V[a, b]$. It will suffice to show that $(\Omega, \infty \oplus 0, [t, b])$ is positive definite for all $t \in [a, b)$. Choose any Λ with property $\oplus$ and order $\Lambda \leqq$ order Ω. By (IV), we have "$\int_a^b \Omega$ positive definite" implies "$(\Omega, \infty, [a, b])$ positive definite" implies "$\mathrm{nul}(\Omega, \infty, [t, b]) = 0$ for $t \in [a, b)$" implies "$(\Omega, \infty, [t, b])$ positive definite." It remains to show that $\int_t^b \Omega$ is positive definite on solutions in $O^{(\nu)}(t)$, i.e., that $c(0, t)$ is positive definite. Now "$\int_a^b \Omega$ positive definite" implies "$c(0, a)$ positive definite" and since $c(0, t) \cap \infty = \{0\}$ for $t \in [a, b)$, the result follows by (II) of Proposition 4.7.

(VI). *It suffices to prove* (A). Since (B) is independent of Λ, we can choose Λ to satisfy the conditions of (I), so (B) = RHS(A). If (A) holds, then (B) = LHS(A) regardless of Λ.

(VII). (A) *holds in the case where order* $\Lambda \leqq$ *order* Ω *and* $\mathrm{nul}(\Omega, \beta, [a, b]) = \mathrm{nul}(\Omega, \infty, [a, b]) = 0$. By (IV) and (II), we have $\mathrm{ix}(\Omega, \infty, [a, b]) = [\alpha(\lambda) : \Gamma_\infty]_{0 \leqq \lambda \leqq \lambda^+}$ for λ^+ sufficiently large. Using Proposition 2.6 and (II), the statement to be proved becomes

$$(3.3) \qquad [\alpha(\lambda) : \Gamma_\beta]_{0 \leqq \lambda \leqq \lambda^+} - [\alpha(\lambda) : \Gamma_\infty]_{0 \leqq \lambda \leqq \lambda^+} = \mathrm{ix}(\alpha(0) - \beta)$$

for all large λ^+. Since we have by (B) of Proposition 4.8 that $[c : \Gamma_\beta] - [c : \Gamma_\infty] = 0$ for any closed curve c, it follows that LHS(3.3) depends only on $\alpha(0)$ and $\alpha(\lambda^+)$. Now let δ be a positive definite hermitian form on $E^\nu \oplus E^\nu$. Then it is easily seen that RHS(3.3) is equal to $[\alpha(0) + s\delta : \Gamma_\beta]_{0 \leqq s \leqq s^+}$ for all s^+ large. Thus we want to show that $\alpha(0) + s^+\delta$ and $\alpha(\lambda^+)$ are in the same component of $\mathrm{Herm}(E^\nu \oplus E^\nu) - \Gamma_\beta - \Gamma_\infty$ for s^+, λ^+ large. From (IV) of Proposition 4.7, this means we must show that $\mathrm{ix}(\alpha(\lambda^+) - \beta \mid \mathrm{domain}\, \beta) = 0$ for λ^+ large. This is done by letting δ be a derivative dependent hermitian form of lesser order such that $\beta - [-\delta(a)] \oplus [\delta(b)]$ is positive definite on domain β and using Lemma 3.2 to find a λ^+ for which $\int_a^b (\Omega + d\delta + \lambda^+\Lambda)$ is positive definite.

(VIII). $\mathrm{ix}(\Omega, \beta, [a, b])$ *is finite.* Choose Λ with order $\Lambda \leqq$ order Ω and $\int_a^b \Lambda$ positive definite. Then $\alpha(\lambda)$ is also defined for negative values of λ near zero. Since $\lambda \to \alpha(\lambda)$ is a $\oplus$-curve, the conditions of (VII) are satisfied by $(\Omega - \varepsilon\Lambda, \beta, [a, b])$ for all small $\varepsilon > 0$. Then $\mathrm{ix}(\Omega, \beta, [a, b]) \leqq \mathrm{ix}(\Omega - \varepsilon\Lambda, \beta, [a, b]) = \mathrm{ix}(\Omega - \varepsilon\Lambda, \infty, [a, b]) + \mathrm{ix}(\alpha(-\varepsilon) - \beta) < \infty$.

(IX). (A) *is true.* By (VII) we have for all but a discrete set of $\varepsilon > 0$ that

$$(\mathrm{A}_\varepsilon) \qquad \mathrm{ix}(\Omega + \varepsilon\Lambda, \beta, [a, b]) = \textstyle\sum_{\lambda > \varepsilon} \mathrm{nul}(\Omega + \lambda\Lambda, \beta, [a, b]) \,.$$

From (VIII) it follows easily that LHS(A_ε) $\leqq$ LHS(A) for all ε sufficiently small. Hence choosing a small ε for which (A_ε) is true we have

$$\text{LHS}(A) = \text{LHS}(A_\varepsilon) = \text{RHS}(A_\varepsilon) \leqq \text{RHS}(A)$$

and, combining this with (III), (A) is proved.

We conclude with the proof of Lemma 3.2. Throughout the remainder of this paragraph Ω and Λ are Sturm forms on $V[a, b]$ and Λ has property $\oplus$. We want to show that $\Omega + \lambda\Lambda$ has property $\oplus$ for λ sufficiently large.

LEMMA 3.3. *Ω has property $\oplus$ if and only if:*

(1) *There is some $\gamma \in \text{Quad}(E^\vee)$ such that for all $t \in [a, b]$ we have $c(t) \cap \gamma = \{0\}$ where, as in* (I), *we define $c(t) = A^\#(b)A^\#(t)^{-1}[\infty]$.*

(2) *$\alpha(\Omega, [a, b])$ is positive definite.*

COROLLARY. *Lemma 3.2 is true in the case where order $\Lambda \geqq$ order Ω.*

DEDUCTION OF COROLLARY. Let $\Omega_s = s\Omega + \Lambda$. Then $\alpha(\Omega_s, [a, b])$ and $c_s(t)$ depend continuously on s. Since properties (1) and (2) remain unchanged under small variation, it follows that Ω_s and hence $\Omega + s^{-1}\Lambda$ has property $\oplus$ for small positive s. (Note: if order $\Lambda <$ order Ω, then the order of Ω_s jumps at $s = 0$ and this argument fails.)

PROOF OF LEMMA 3.3. Since $t \to \Phi(t) = A^\#(t)^{-1}[\infty]$ is a $\oplus$-curve, it is easily seen by (II) of Proposition 4.7 that (1) is fulfilled if and only if $\Phi(t_0) \cap \Phi(t_1) = \{0\}$ whenever $t_0, t_1 \in [a, b]$ and $t_0 \neq t_1$. Thus if (1) holds, $(\Omega, \infty, [a, b])$ is non-degenerate and, by Proposition 2.4, it has index zero. Then (2) and Proposition 2.6 show that $\int_a^b \Omega$ is positive definite. Then since $c(t) \cap \infty = \{0\}$ for $t \in [a, b)$, the concluding argument of (V) shows that Ω has property $\oplus$. That property $\oplus$ implies (1) and (2) is immediate.

LEMMA 3.4. *If Λ is of order zero, then there is an $\varepsilon > 0$ with the property that, for any subinterval $[t_0, t_1] \subset [a, b]$ of length $<\varepsilon$, there is a λ such that $\int_{t_0}^{t_1}(\Omega + \lambda\Lambda)$ has property $\oplus$ on $V[t_0, t_1]$.*

COROLLARY. *Lemma 3.2 is true in the case where order Λ is zero.*

DEDUCTION OF COROLLARY. Subdivide $a = t_0 < t_1 < \cdots < t_n = b$ with $t_i - t_{i-1} < \varepsilon$, and choose λ_i such that $\Omega + \lambda_i\Lambda$ has property $\oplus$ on $V[t_{i-1}, t_i]$. Setting $\lambda = \max\{\lambda_i\}$, it is immediate that $\Omega + \lambda\Lambda$ has property $\oplus$ on $V[a, b]$.

PROOF OF LEMMA 3.4. By Theorem 2.1, there is an ε with the property that $t_1 - t_0 < \varepsilon$ implies $\Phi(t) \cap \Phi(t_1) = \{0\}$ for $t \in [t_0, t_1)$. Since this implies (1) of Lemma 3.3 for $\Omega + \lambda\Lambda$ on $V[t_0, t_1]$ (all λ) it will suffice to prove that there is an ε with the property that $t_1 - t_0 < \varepsilon$ implies $\int_{t_0}^{t_1}(\Omega + \lambda\Lambda)$ is positive definite for λ sufficiently large. This is done by the following construction:

For $s, t \in [a, b]$ with $s \neq t$, let $|[s, t]|$ denote $[s, t]$ if $s < t$ and $[t, s]$ if $t < s$. Let $m(s, t) : |[s, t]| \to [-1, 1]$ be the map $r \to (2r - t - s)/(t - s)$. Then $m(s, t)$ induces $V[-1, 1] \to V|[s, t]|$ which, in turn, induces a map from Sturm forms on $V|[s, t]|$ to Sturm forms on $V[-1, 1]$. We denote by $m(s, t)^\flat[\Omega]$ the Sturm form obtained by restricting Ω to $V|[s,t]|$ and applying $m(s,t)$. A simple computation shows that, if $\omega_{ij}(r)$: $E \to E^*$ is the "coordinate representation" of Ω as in (1.1), then the coordinate representation of $m(s,t)^\flat[\Omega]$ is given by

$$\{m(s, t)^\flat[\Omega]\}_{ij}(q) = \left(\frac{2}{t - s}\right)^{j+i} \omega_{ij}\left(\frac{(t - s)q + t + s}{2}\right) \tag{3.4}$$

where $q \in [-1, 1]$. Set $\Omega^\flat(s, t) = m(s, t)^\flat[(t - s)^{2\nu}\Omega + \Lambda]$ where $\nu =$ degree Ω. For the case $s = t$, we define $\Omega^\flat(t, t)$ to be the constant coefficients form whose leading coefficient is $2^{2\nu}\omega_{\nu\nu}(t)$ and whose 0^{th} order term is $\Lambda(t)$ (with all other terms zero). Then (3.4) shows that $\Omega^\flat(s, t)$ depends smoothly on $(s, t) \in [a, b] \times [a, b]$. By inspection $\Omega^\flat(t, t)$ has property $\oplus$, hence by continuity and the preceding lemma, it follows that $\Omega^\flat(s, t)$ has property $\oplus$ on some neighborhood of the diagonal in $[a, b] \times [a, b]$, hence there is an ε such that $[t_0, t_1] \subset [a, b]$, and $t_1 - t_0 < \varepsilon$ implies that $\int_{t_0}^{t_1}(\Omega + (t_1 - t_0)^{-2\nu}\Lambda)$ is positive definite on $V[t_0, t_1]$ as was to be shown.

PROOF OF LEMMA 3.2. The case $0 <$ order $\Lambda <$ order Ω remains. Choose Λ_0 of order zero. From Lemma 3.3, there is an $\varepsilon > 0$ such that $-\varepsilon\Lambda_0 + \Lambda$ has property $\oplus$. From Lemma 3.4, there is a K such that $\Omega + K\Lambda_0$ has property $\oplus$. Hence $\Omega + (K/\varepsilon)\Lambda$ has property $\oplus$, and the lemma is proved.

4. U-manifolds

4.1. *Definitions and examples.*

DEFINITION. By a U-*manifold* we mean a set $\text{U}(E, \psi)$ obtained from an even dimensional complex vector space E and a non-degenerate hermitian form ψ on E of signature zero [i.e., $\text{nul}(\psi) = 0, \text{ix}(\psi) = \text{ix}(-\psi)$] by setting $\text{U}(E, \psi) = \{$all subspaces $P \subset E$ of dimension $(\dim E)/2$ with the property that $\psi \mid P \equiv 0\}$.

REMARKS. 1. A non-degenerate ψ has signature zero if and only if there is a P of half the dimension for which $\psi \mid P = 0$; i.e., if and only if $\text{U}(E, \psi)$ is non-empty. For the ψ of 1.3, the subspace $\Phi(t) = S \cap O^{(\nu)}(t)$ shows that the signature is zero.

2. Let G be the Grassmann manifold of all subspaces $P \subset E$ of half the dimension of E. Then $\text{U}(E, \psi)$ can also be described as the fixed point set of the involution $P \to P^\perp$ (rel. ψ) of G. In this way $\text{U}(E, \psi)$ becomes a topological space, and in fact a manifold.

3. Since two non-degenerate hermitian forms on E are equivalent if and only if they have the same signature, it follows that $\mathrm{U}(E, \psi)$ is determined, up to isomorphism, by the dimension of E.

Example 1. Let F be a complex vector space with a given hermitian norm $\| \ \|^2$. Let ψ be the hermitian form on $F \oplus F$ defined by $\psi[(x, y)] = \|x\|^2 - \|y\|^2$. Then $\mathrm{U}(F \oplus F, \psi)$ is the set of all graphs of norm-preserving linear maps $F \to F$.

REMARK. This example shows that as a topological space $\mathrm{U}(E, \psi)$ is homeomorphic to the unitary group $\mathrm{U}(n)$ where $2n = \dim E$.

Example 2. $E = F \oplus F^*$, where F is a finite dimensional complex vector space, and ψ is defined by $\psi[(x, y)] = \mathrm{Im}\langle x, y\rangle$. Then $\mathrm{U}(E, \psi)$ has the following interpretation:

DEFINITION. By a *hermitian pair* on a complex vector space F, we mean a pair (Σ, β) where Σ is a subspace of F and β is a hermitian form on Σ.

PROPOSITION 4.1. *The elements of* $\mathrm{U}(F \oplus F^*, \psi)$ *of Example* 2 *correspond one-one with hermitian pairs* (Σ, β) *on* F *by the rule*

$$(\Sigma, \beta) \leftrightarrow \{(x, y) \in F \oplus F^* : x \in \Sigma \text{ and } \langle z, y\rangle = \beta[z, x] \text{ for all } z \in \Sigma\}\ .$$

PROOF. It is immediately seen that the RHS is a subspace on which ψ is zero. That its dimension is $\dim F$ is seen from the fact that its projection on the first coordinate has image $\Sigma \oplus 0$ and kernel $0 \oplus \Sigma^{\perp}$. Conversely if $P \in \mathrm{U}(F \oplus F^*, \psi)$, set $\Sigma_p =$ projection of P on first coordinate, and define $\beta_p[x] = \langle x, y\rangle$ where $(x, y) \in P$. Then β_p is real-valued and we need only show it is well-defined. But if $(x, y), (x, y') \in P$, we have $(x, y + i(y - y')) \in P$ hence $\mathrm{Im}\langle x, y + i(y - y')\rangle = 0$ from which $\mathrm{Re}\langle x, y\rangle = \mathrm{Re}\langle x, y'\rangle$.

DEFINITION. The set of all hermitian pairs on F, considered as the U-manifold of Example 2, will be denoted $\mathrm{Herm}(F)$. The element $F \oplus 0$, i.e., the pair $(F, 0)$, will be denoted by $\mathbf{0} \in \mathrm{Herm}(F)$. The element $0 \oplus F^*$, i.e., the pair $(\{0\}, 0)$, will be denoted by $\boldsymbol{\infty} \in \mathrm{Herm}(F)$.

REMARKS. 1. If $F = \mathbf{C}$, then the Grassman manifold G is the complex projective line (i.e., the Riemann sphere) the involution is complex conjugation, and $\mathrm{Herm}(F)$ is the real axis plus the point at ∞.

2. The correspondence between Examples 1 and 2 is given by the Cayley transform.

4.2. *The structure of* U-*manifolds*. Let $U = \mathrm{U}(E, \psi)$. By an automorphism of U we mean a map $U \to U$ induced by an automorphism of E which leaves ψ invariant. From Example 1, it is easily seen that the group of automorphisms acts transitively on U. From Example 2, it is easy to deduce the following proposition:

DEFINITION. For $P \in U$, set $\Gamma_P = \{Q \in U : P \cap Q \neq 0\}$. [We mean, of course, $P \cap Q \neq \{0\}$ but will often overlook this detail for the sake of simplicity.]

PROPOSITION 4.2. *The group of all automorphisms of U which leave P fixed acts transitively on $U - \Gamma_P$ (set theoretic—).*

PROOF. Take $P = \infty \in \mathrm{Herm}(F)$.

Otherwise stated, for any $P_1, P_2, Q_1, Q_2 \in U$ with $P_i \cap Q_i = 0$ there is an automorphism of U carrying $P_1 \to P_2$, $Q_1 \to Q_2$ so any pair of disjoint planes is like any other. However, three disjoint planes give rise to the following non-trivial invariant:

DEFINITION. Given $P, Q, R \in U$ with $P \cap R = 0$, define a hermitian form $\alpha(P, Q, R)$ on Q as follows: Let j: $E \to E$ be the unique map which is the identity on P and multiplication by $-i$ on R. Then set $\alpha(P, Q, R)[q] = \psi[jq]$ for $q \in Q$.

PROPOSITION 4.3. *Given $P, R \in U$ with $P \cap R = 0$. Then for $Q \in U$ we have:*

(1) *$\alpha(P, Q, R)$ is non-degenerate if and only if $P \cap Q = R \cap Q = 0$*, i.e., *if and only if P, Q, R are mutually disjoint.*

(2) *$Q \to \mathrm{ix}\,\alpha(P, Q, R)$ is a continuous integer-valued function on $U - \Gamma_P - \Gamma_R$ which distinguishes components. In particular, $U - \Gamma_P - \Gamma_R$ has* $(\dim E/2) + 1$ *components.*

PROOF. We can take $U = \mathrm{Herm}(F)$, $P = 0$, $R = \infty$. It is easily checked that $x \in P \cap Q$ or $x \in R \cap Q$ implies x is in the nullspace of $\alpha(P, Q, R)$. Suppose then that $P \cap Q = R \cap Q = 0$. Then Q is of the form $\{(x, \alpha_Q x)\}$ where α_Q: $F \to F^*$ is self dual and non-singular. Then $\alpha(P, Q, R)$ is the map $(x, \alpha_Q x) \to \langle x, \alpha_Q x\rangle$ which is non-degenerate and (1) is proved. (2) is simply the statement that the space of all non-degenerate hermitian forms on an n-dimensional space F has $(n + 1)$ components which are distinguished by the index.

Near P this labelling of the components of $U - \Gamma_P - \Gamma_R$ is independent of R. Specifically:

PROPOSITION 4.4. *Let $P, R_1, R_2 \in U$ be such that $P \cap R_1 = P \cap R_2 = 0$. Then there is a neighborhood N of P, $P \in N \subset U - \Gamma_{R_1} - \Gamma_{R_2}$, such that for all $Q \in N \cap (U - \Gamma_P)$, we have $\mathrm{ix}\,\alpha(P, Q, R_1) = \mathrm{ix}\,\alpha(P, Q, R_2)$.*

PROOF. Join R_1 to R_2 by a curve $R(t)$ in $U - \Gamma_P$. Then the union of the

$\Gamma_{R(t)}$ is a closed set not containing P. Let N be its complement. Then for $Q \in N \cap (U - \Gamma_P)$ the integer $\mathrm{ix}\,\alpha(P, Q, R(t))$ depends continuously on t, hence is constant.

DEFINITION. A curve Q: $[a, b] \to U$ is called a $\oplus$-*curve* if, for any $t_0 \in [a, b)$ and any $R \in U$ with $Q(t_0) \cap R = 0$, there is an $\varepsilon > 0$ such that $\alpha(Q(t_0), Q(t), R)$ is positive definite for all $t \in (t_0, t_0 + \varepsilon)$.

PROPOSITION 4.5. *Let* Q: $[a, b] \to U$ *be a* $\oplus$-*curve. Then for any* $P \in U$, *the curve* Q *intersects* Γ_P *in a discrete set of points. Moreover, if there is an* $R \in U$ *such that the curve* Q *does not intersect* Γ_R, *then*

$$\begin{aligned}\mathrm{ix}\,\alpha(P, Q(a), R) &= \mathrm{ix}\,\alpha(P, Q(b), R) + \textstyle\sum_{a<t\leq b} \mathrm{nul}\,\alpha(P, Q(t), R)\\ &= \mathrm{ix}\,\alpha(P, Q(b), R) + \textstyle\sum_{a<t\leq b} \dim[P \cap Q(t)]\,.\end{aligned}$$

PROOF. The proposition reduces easily to the following statement: Let $t \to \beta(t)$ be a curve of hermitian forms on F with the property that $\beta(t_1) - \beta(t_0)$ is positive definite whenever $t_1 > t_0$. Then $\beta(t)$ is non-degenerate except at a discrete set of t's, and for small $\varepsilon > 0$ we have $\mathrm{ix}\beta(t_0 + \varepsilon) = \mathrm{ix}\beta(t_0)$, $\mathrm{ix}\beta(t_0 - \varepsilon) = \mathrm{ix}\beta(t_0) + \mathrm{nul}\beta(t_0)$. This is seen by taking a decomposition $F = F_+ \oplus F_0 \oplus F_-$ such that F_0 is the nullspace of $\beta(t_0)$, while F_+ (resp. F_-) is a subspace on which $\beta(t_0)$ is positive (resp. negative) definite. Then $\beta(t_0 + \varepsilon)$ is positive definite on $F_+ \oplus F_0$ and negative definite on F_-, from which it follows that $\beta(t_0 + \varepsilon)$ is non-degenerate with index equal to $\dim F_-$. Similarly $\beta(t_0 - \varepsilon)$ is non-degenerate with index $\dim(F_- \oplus F_0)$.

PROPOSITION 4.6. *Let* Q: $[a, b] \to U$ *be a differentiable curve, and let* $t_0 \in [a, b]$. *Then the hermitian form* $(d/dt)\,|_{t_0}\,(P, Q(t_0), Q(t))$ (*the derivative of a curve of hermitian forms on* $Q(t_0)$ *is a hermitian form on* $Q(t_0)$), *defined for any* $P \in U - \Gamma_Q(t_0)$, *is independent of* P.

DEFINITION. The above form on $Q(t_0)$ will be called the *tangent vector* to the curve Q at $Q(t_0)$.

REMARKS. 1. It is easily seen that the tangent vector is, as the name implies, a complete characterization of the first order properties of the curve Q at $Q(t_0)$.

2. It is easily proved that a curve all of whose tangent vectors are positive definite is a $\oplus$-curve, but not conversely.

PROOF OF PROPOSITION. Let $q \in Q(t_0)$. For any P complementary to $Q(t_0)$ we have a unique decomposition $q = p(t) + q(t)$ where $p(t) \in P$, $q(t) \in Q(t)$ (for t near t_0). Then computation gives $\alpha(P, Q(t_0), Q(t))[q] = 2\,\mathrm{Im}\,\psi[q(t), q]$. Now let $t \to \bar{q}(t)$ be any differentiable curve in E with $\bar{q}(t_0) = q$ and $\bar{q}(t) \in Q(t)$.

Then differentiating the relation $\psi[q(t), \bar{q}(t)] \equiv 0$ we have

$$(4.1) \qquad \frac{d}{dt}\bigg|_{t_0} 2\,\mathrm{Im}\,\psi[q(t), q] = \frac{d}{dt}\bigg|_{t_0} 2\,\mathrm{Im}\,\psi[\bar{q}(t), q]$$

which proves not only the proposition but also:

COROLLARY TO PROOF. *The tangent vector to $t \to Q(t)$ at $Q(t_0)$ evaluated on $q \in Q(t_0)$ is equal to* LHS (4.1) *where $t \to q(t)$ is any differentiable curve with $q(t) \in Q(t)$, $q(t_0) = q$.*

We conclude this section with a catch-all proposition which is simply a list of properties needed elsewhere. The proofs are elementary.

PROPOSITION 4.7. (I). *A curve $t \to Q(t) \in U$ is a $\oplus$-curve if and only if it has the property that, for t_0 in its domain and $P \in U - \Gamma_{Q(t_0)}$, there is an $\varepsilon > 0$ such that $\alpha(P, Q(t), Q(t_0))$ is positive definite for $t \in (t_0 - \varepsilon, t_0)$.*

(II). *If Q: $[a, b] \to U$ is a $\oplus$-curve, if $\alpha(P, Q(a), R)$ is positive definite, and if $Q(t) \cap R = 0$ for $t \in [a, b]$, then $Q(t) \cap P = 0$ for $t \in [a, b]$.*

(III). *A curve Q: $[a, b] \to U$ will be called discretely self-intersecting if there is an $\varepsilon > 0$ for which $Q(t_0) \cap Q(t_1) = 0$ whenever $0 < |t_0 - t_1| < \varepsilon$. Then a differentiable curve Q is a $\oplus$-curve if and only if*

(i) *it is discretely self-intersecting, and*

(ii) *its tangent vector at each point is positive semi-definite.*

(IV). *Given $P_1, P_2, Q_1, Q_2 \in U$, there is an automorphism of U which carries $P_1 \to P_2$ and $Q_1 \to Q_2$ if and only if* $\dim[P_1 \cap Q_1] = \dim[P_2 \cap Q_2]$. *In the case where $U = \mathrm{Herm}(F)$, $Q = \infty$ and $P = (\Sigma, \beta)$, a hermitian form $\gamma \in \mathrm{Herm}(F) - \Gamma_\infty$ is in $\Gamma_{(\Sigma,\beta)}$ if and only if $(\gamma - \beta \,|\, \Sigma)$ is degenerate, from which it follows that $\mathrm{Herm}(F) - \Gamma_\infty - \Gamma_{(\Sigma,\beta)}$ has $(\dim \Sigma + 1)$ components which are distinguished by the function $\gamma \to \mathrm{ix}(\gamma - \beta \,|\, \Sigma)$.*

4.3. *Intersection theory.* The intersection theory is specified by the following axioms:

Fix $Q \in U$.

Axiom 1. An integer, denoted $[P(t) : \Gamma_Q]_{a \le t \le b}$ and called the total intersection of P with Γ_Q, is assigned to each (continuous) curve P: $[a, b] \to U$ for which $P(a), P(b) \notin \Gamma_Q$.

Axiom 2. If $(s, t) \to P_s(t)$ is a map of $[0, 1] \times [a, b] \to U$ with $P_s(a)$, $P_s(b) \notin \Gamma_Q$ for $s \in [0, 1]$, then $[P_s(t) : \Gamma_Q]_{a \le t \le b}$ is independent of s.

Axiom 3. If P: $[a, c] \to U$ and $b \in (a, c)$ are given, and if $P(a)$, $P(b)$, $P(c) \notin \Gamma_Q$, then

$$[P(t) : \Gamma_Q]_{a \le t \le c} = [P(t) : \Gamma_Q]_{a \le t \le b} + [P(t) : \Gamma_Q]_{b \le t \le c} \,.$$

Axiom 4. If $P : [a, b] \to U$ is a $\oplus$-curve such that $P(t) \in \Gamma_Q$ for only one

$t_0 \in (a, b)$, and if $\dim [P(t_0) \cap Q] = 1$, then $[P(t) : \Gamma_Q]_{a \leqq t \leqq b} = 1$.

PROPOSITION 4.8. *Axioms* 1–4 *uniquely determine* $[P(t) : \Gamma_Q]_{a \leqq t \leqq b}$. *We have the further properties that:*

(A). *If P is a $\oplus$-curve, then* $[P(t) : \Gamma_Q]_{a \leqq t \leqq b} = \sum_{a<t<b} \dim [P(t) \cap Q]$.

(B). *If* $P(a) = P(b)$, *then* $[P(t) : \Gamma_Q]_{a \leqq t \leqq b}$ *is independent of Q (as long as it is defined;* i.e., $P(a) \notin \Gamma_Q$).

PROOF. Axioms 1–3 say that $[P : \Gamma_Q]$ is a homeomorphism of the relative homotopy group $\pi_1(U, U - \Gamma_Q)$ into the integers $\mathbf{Z}$. Since U is homeomorphic to the unitary group and $U - \Gamma_Q$ is a cell, it follows that $\pi_1(U, U - \Gamma_Q) = \mathbf{Z}$ and Axiom 4 will determine the homeomorphism uniquely. It remains only to show that Axiom 4 can be fulfilled; i.e., that any two curves of the type it describes are in the same class of $\pi_1(U, U - \Gamma_Q)$ and that this class is a generator. To prove this, let P_1, P_2 be two such curves. Using Axiom 2 to shrink their domains if necessary we choose a representation of U as Herm(F) with $Q = \mathbf{0}$ in which P_1 and P_2 lie in Herm$(F) - \Gamma_\infty$. Choosing a basis in F we see from Proposition 4.5 that, modulo deformations of P_i in $U - \Gamma_\infty$ keeping end points in $U - \Gamma_0 - \Gamma_\infty$, we can assume that the P_i are of the form

$$t \to \operatorname{diag}(-1, -1, \cdots, -1, t, 1, 1, \cdots, 1) \qquad t \in [-1, 1] \tag{4.2}$$

there being $\dim F$ such curves, one from each component of $U - \Gamma_0 - \Gamma_\infty$ up to the next component. Taking the Cayley transform they become

$$\theta \to \operatorname{diag}(-i, -i, \cdots, -i, e^{i\theta}, i, i, \cdots, i) \qquad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Now these curves are homotopic mod $(U - \Gamma_Q)$ (Q is the identity matrix) to the closed curves $\operatorname{diag}(-1, \cdots, -1, e^{i\theta}, -1, \cdots, -1)$ for $\theta \in [-\pi, \pi]$ which are all representatives of the same generator of $\pi_1(U)$. Thus P_1 and P_2 are in the same class of $\pi_1(U, U - \Gamma_Q)$, and that class generates.

To prove Property (A), we can treat each of the values t_0 for which $P(t_0) \cap Q \neq 0$ individually. Choose a representation of U as Herm(F) in which $Q = \mathbf{0}$ and $P(t_0) \notin \Gamma_\infty$. Then from Proposition 4.5, we see that P is homotopic mod $(U - \Gamma_Q)$ to k of the curves (4.2) laid end to end, where $k = \dim [P(t_0) \cap Q]$. Then (A) follows from the proof and *Axiom* 3.

Property (B) follows by observing that the intersection theory for Q was shown to arise from a choice of generator [the class of $\operatorname{diag}(-1, \cdots, -1, e^{i\theta}, -1, \cdots, -1)$] of $\pi_1(U)$ followed by $\pi_1(U) \approx \pi_1(U, U - \Gamma_Q)$. Hence we need only show that the choice of this generator of $\pi_1(U)$ is independent of Q. This follows from property (A), which shows that a closed $\oplus$-curve must have positive total intersection with Γ_Q regardless of Q.

5. Eigenfunction expansions

The relevance of the generalized Sturm theorem of the existence of eigenfunction expansions is seen from the following general observation:

DEFINITION. Let Q and R be hermitian forms on V. An eigenvector of of the pair (Q, R) is a non-zero vector v for which there is a complement $W \subset V$ with the property that $V = \{cv\} \oplus W$ splits both Q and R. (If either $Q[v]$ or $R[v]$ is non-zero, then W is unique).

PROPOSITION 5.1. *Let Q, R be hermitian forms on V with R positive definite. For any real λ, let W_λ be the linear span of all eigenvectors v of (Q, R) for which $(Q + \lambda R)[v] < 0$. We assume:*

(1) *W_λ is finite dimensional, and*

(2) *$Q + \lambda R$ is positive semi-definite on $W_\lambda^\perp$* (rel. R).

For a given $x \in V$, let x_λ be the orthogonal projection (rel. R) *of x on W_λ. Then $\lim_{\lambda\to-\infty} R[x - x_\lambda] = 0$.*

PROOF. We will call $W \subset V$ a *splitting space* of (Q, R) if it has a complement W' such that $W \oplus W'$ splits Q and R. (Thus an eigenvector is a one-dimensional splitting space.) Since the span and intersection of two splitting spaces are again splitting spaces, it follows that the W_λ are splitting spaces of (Q, R). Without loss of generality we can assume that Q is positive definite because (1) and (2) imply that $Q + KR$ is positive definite for K large, and changing Q to $Q + KR$ does not change the conclusion. Now by (2) we have $(Q + \lambda R)[x - x_\lambda] \geqq 0$, so $(-\lambda)^{-1}Q[x - x_\lambda] \geqq R[x - x_\lambda]$, and it suffices to show that $Q[x - x_\lambda]$ is bounded as $\lambda \to -\infty$, which follows from

$$Q[x] = Q[x_\lambda] + Q[x - x_\lambda] \geqq Q[x - x_\lambda] .$$

REMARK. Under the assumption that R is positive definite, we have v an eigenvector if and only if $\{cv\}^\perp(\text{rel. } R) \subset \{cv\}^\perp(\text{rel. } Q)$ if and only if $(Q + \lambda R)[w, v] = 0$ for all $w \in V$ where $\lambda = -Q[v]/R[v]$. Hence the present definition of eigenvector reduces to the usual one.

COROLLARY. *Let Ω, Λ be Sturm forms on $V[a, b]$ with order $\Lambda <$ order Ω and $\int_a^b \Lambda$ positive definite, and let $\beta \in \operatorname{Herm}(E^\nu \oplus E^\nu)$. Choose a sequence $y_1, y_2, \cdots \in V[a, b]$ such that:*

(1) *y_i is in the nullspace of $(\Omega + \lambda_i\Lambda, \beta, [a, b])$ for some λ_i.*

(2) $\lambda_i \geqq \lambda_{i+1}$.

(3) *For any real λ, the y_i which are in nullspace of $(\Omega + \lambda\Lambda, \beta, [a, b])$ are a basis of this nullspace and are orthonormal with respect to $\int_a^b \Lambda$.*

Given $x \in$ domain $(\Omega, \beta, [a, b])$, set

$$c_i = \int_a^b \Lambda[x, y_i] .$$

Then

$$\lim_{n\to\infty}\int_a^b \Lambda[x - \textstyle\sum_{i=1}^n c_i y_i] = 0 .$$

PROOF. In the notation of the proposition, let $V =$ domain $(\Omega, \beta, [a, b])$, $Q = (\Omega, \beta, [a, b])$, $R = \int_a^b \Lambda$. Then v is an eigenvector of (Q, R) if and only if it is a non-zero element of the nullspace of $(\Omega + \lambda\Lambda, \beta, [a, b])$ where $\lambda = -Q[v]/R[v]$. Thus W_λ is a direct sum of nullspaces, and by Theorem 3.1, the sum contains a finite number of finite-dimensional terms. More importantly, Theorem 3.1 says that the dimension of W_λ is equal to the index of $(Q + \lambda R)$ so, since $Q + \lambda R$ is negative definite on W_λ, it follows [W_λ is a splitting space of (Q, R)] that (2) is satisfied. Thus the conclusion of Proposition 5.1 holds. The y's and c's simply constitute a computation of x_λ.

A more familiar statement would be:

COROLLARY (Sturm-Liouville theorem). *Let $(\Omega, \beta, [a, b])$ be given on $V[a, b]$. Take Λ to be $\Lambda(t)[x] = \| x(t) \|^2$ where $\| \ \|$ is a norm on E. Let H be the Hilbert space obtained by taking the completion of domain $(\Omega, \beta, [a, b])$ relative to the L^2 norm $\int_a^b \Lambda$. Then the functions y_i defined by the relevant eigenvalue problem (see* § 6) *and by* (1)–(3) *above, are a complete orthonormal set in H.*

PROOF. The assertion is that linear combinations of the y_i are dense in H. It follows from the proposition that they are dense in domain $(\Omega, \beta, [a, b])$ hence they are dense in H.

6. The classical formulation

PROPOSITION 6.1. *Let Ω be a non-degenerate derivative dependent hermitian form of order ν on $V[a, b]$, and let $\langle \, , \rangle$ be a hermitian norm on the image space E of $V[a, b]$. Then the formula* (1.2) *associates to Ω a formally self-adjoint linear differential operator L of order 2ν. Moreover, there is a natural one-one correspondence* Π: Herm$(E^\nu \oplus E^\nu) \to$ {*self-adjoint boundary conditions for L on $[a, b]$*} *such that x is in the nullspace of $(\Omega, \beta, [a, b])$ if and only if $Lx \equiv 0$ and x satisfies the boundary condition* $\Pi(\beta)$. *Any self-adjoint boundary value problem* (L, Π) *in which L is formally self-adjoint and of even order arises in this way.*

PROOF. To see that the L of (1.2) is formally self-adjoint, use the identity $\Omega[x, y] = \overline{\Omega[y, x]}$ to obtain

$$\langle x, Ly\rangle - \langle Lx, y\rangle = \frac{d}{dt}\{\langle A\mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{x}, A\mathbf{y}\rangle\} . \tag{6.1}$$

Now for any L there are unique L^* and $B(t)$: $E^{2\nu} \times E^{2\nu} \to C$ (sesqui-linear) such that

$$(6.2) \qquad \langle x, L\mathbf{y}\rangle - \langle L^*\mathbf{x}, y\rangle = \frac{d}{dt} B(t)[\mathbf{x}, \mathbf{y}]$$

by integration by parts. Thus (6.1) gives $L = L^*$ and $B(t)[\mathbf{x}, \mathbf{y}] = \langle A\mathbf{x}, \mathbf{y}\rangle - \langle \mathbf{x}, A\mathbf{y}\rangle$. Now when L is formally self-adjoint, i.e., $L = L^*$, a self-adjoint boundary condition for L on $[a, b]$ is by definition a subspace $\Pi \subset E^{4\nu}$ such that

$$(6.3) \qquad -B(a)[(v_1, v_2), (v_1, v_2)] + B(b)[(v_3, v_4), (v_3, v_4)] = 0$$

for all $(v_1, v_2, v_3, v_4) \in \Pi$. Let SA be the set of all such Π. Since B is (by uniqueness) skew hermitian, the real part of (6.3) is zero. Its imaginary part, on the other hand, is

$$-2\,\mathrm{Im}\langle v_1, A(a)(v_1, v_2)\rangle + 2\,\mathrm{Im}\langle v_3, A(b)(v_3, v_4)\rangle .$$

Define $\mathfrak{A}$: $E^{4\nu} \to E^{\nu} \oplus E^{\nu *} \oplus E^{\nu} \oplus E^{\nu *}$ by $\mathfrak{A}(v_1, v_2, v_3, v_4) = (v_1, A(a)(v_1, v_2), v_3, A(b)(v_3, v_4))$. In terms of the notation of Definition 2.7, we then have $SA = \{\Pi \subset E^4\text{: } \mathfrak{A}[\Pi] \in \mathrm{U}(F, \Psi)\}$. Hence $\mathfrak{A}^{-1}$ (which exists for the same reason that $A^{\#}$ is non-singular) induces $\mathrm{Herm}(E^{\nu} \oplus E^{\nu}) \to SA$. Take this map as the Π of the proposition. If $x \in V[a, b]$, then it is easily seen from Proposition 4.1 that $Lx \equiv 0$ and $(\mathbf{x}^{(2\nu)}(a), \mathbf{x}^{(2\nu)}(b)) \in \Pi(\beta)$ if and only if $\hat{x} \in$ domain β and $\int_a^b \Omega[y, x] = \beta[\hat{x}, \hat{y}]$ for all $y \in V[a, b]$ with $\hat{y} \in$ domain β [where $\hat{x}$ means $(\mathbf{x}^{(\nu)}(a), \mathbf{x}^{(\nu)}(b))$ and similarly $\hat{y}$].

It remains only to show that every (L, Π) arises in this way, and for this, it will suffice to show that every even order self-adjoint L arises from some Ω *via* (1.2). This is done by integrating $\langle x, Ly\rangle$ by parts to put it in the form $\sum \langle x^{(i)}, \omega_{ij} y^{(j)}\rangle$ where $\omega_{ij} \equiv 0$ for $|i - j| > 1$ and $\omega_{i,i+1} + \omega_{i+1,i} \equiv 0$. Then the assumption $L = L^*$ and (6.2) imply, without too much difficulty, that $\omega_{i,j} = \omega_{j,i}$ and that Ω defined by $\Omega[x] = \sum \langle x^{(i)}, \omega_{ij} x^{(j)}\rangle$ gives rise to L.

Remark. We have as an easy corollary that Ω is unique under the assumptions $\omega_{ij} = 0$ for $|i - j| > 1$ and $\omega_{i,i+1} + \omega_{i+1,i} = 0$. In this case the $\omega_{i,i}$ are hermitian and the terms $\omega_{i,i+1}$ and $\omega_{i+1,i}$ are skew hermitian. This yields a well-known canonical form for a self-adjoint L.

7. The Sturm theorems

It has already been pointed out that the Sturm oscillation theorem is the equality (B) = RHS(A) in Theorem 3.1.

COMPARISON THEOREM 7.1. *Let Ω_1, Ω_2 be Sturm forms on $V[a, b]$, and assume that $\Omega_2(t)[x] \geqq \Omega_1(t)[x]$ for all $x \in V[a, b]$ and $t \in [a, b]$. Let*

$\beta \in \mathrm{Herm}(E^{\nu} \oplus E^{\nu})$ *for any positive integer* ν. *Then*

$$\mathrm{ix}(\Omega_2, \beta, [a, b]) \leqq \mathrm{ix}(\Omega_1, \beta, [a, b]) .$$

In this formulation the theorem is obvious.

COROLLARY 1 (Sturm comparison theorem). *Let* $L_i x$ *be defined to be* $-(p_i x')' + r_i x$ *where* p_1, p_2, r_1, r_2 *are differentiable real-valued functions of a real variable* $t \in [a, b]$ *with* $p_2 \geqq p_1 > 0$, $r_2 \geqq r_1$, *and where* x *is a real function of* $t \in [a, b]$. *For* $i = 1, 2$, *let* x_i *be a non-trivial solution of* $L_i x_i = 0$ *for which* $x_i(a) = 0$. *Then* x_1 *has at least as many zeros on* $[a, b]$ *as* x_2 *does.*

PROOF. The number of zeros of x_i is, by (B) of Theorem 3.1, equal to $\mathrm{ix}(\Omega_i, \infty, [a, b])$ where $\Omega_i(t)[x] = p_i x'^2 + r_i x^2$.

COROLLARY 2. *Let* Ω_1, Ω_2 *be as in the theorem, and assume further that* $(\Omega_1, \infty, [a, b])$ *is positive definite and order* $\Omega_1 =$ *order* $\Omega_2 = \nu$. *Let* $\gamma_i \in \mathrm{Herm}(E^{\nu})$ *be the hermitian form such that* $\int_a^b \Omega_i[x] = \gamma_i[\mathbf{x}(b)]$ *for all solutions* x *of* Ω_i *in* $O^{(\nu)}(a)$. *Then* $\gamma_2 \geqq \gamma_1$.

PROOF. γ_i can also be described as the largest hermitian form γ on E^{ν} for which $\mathrm{ix}(\Omega_i, \infty \oplus \gamma, [a, b]) = 0$, and the corollary follows.

The present formulation does not seem to lead to a natural generalization of the Rauch comparison theorem (Rauch, [6, p. 43]), but a substantial part of the proof of that theorem (Lemma 2, p. 43) is a special case of Corollary 2.

SEPARATION THEOREM 7.2. *Let* Ω *be a Sturm form of order* ν *on* $V[a, b]$, *and let* $\gamma_1, \gamma_2 \in \mathrm{Herm}(E^{\nu})$. *Then for any* $[t_0, t_1] \subset [a, b]$, *we have*

$$\sum_{t_0 \leqq t \leqq t_1} \mathrm{nul}(\Omega, \infty \oplus \gamma_1, [t, b]) - \sum_{t_0 \leqq t \leqq t_1} \mathrm{nul}(\Omega, \infty \oplus \gamma_2, [t, b]) \leqq \dim(E^{\nu}) - \dim(\gamma_1 \cap \gamma_2) .$$

PROOF. Let $\bar{\gamma}_i = A^{\#}(b)^{-1}[\gamma_i] \in \cup(S, \psi)$. Then the LHS can be rewritten as

$$[\Phi(t) : \Gamma_{\bar{\gamma}_1}]_{t_0 \leqq t \leqq t_1} - [\Phi(t) : \Gamma_{\bar{\gamma}_2}]_{t_0 \leqq t \leqq t_1}$$

(if $\Phi(t_0)$ or $\Phi(t_1)$ intersects $\bar{\gamma}_1$, or $\bar{\gamma}_2$ extend to a slightly large interval). As in (VII) of the proof of Theorem 3.1, this expression depends only on $\Phi(t_0)$ and $\Phi(t_1)$. The result is then clear from (IV) of Proposition 4.7.

COROLLARY (Sturm separation theorem). *If* x_1, x_2 *are solutions of* $-(px_i')' + rx_i = 0$ *where* x_i, p, r *are differentiable real-valued functions of* $t \in [a, b]$ *with* $p(t) \neq 0$, *then any interval which contains two zeros of* x_1 *must contain a zero of* x_2.

PROOF. Take $\Omega(t)[x] = \pm(px'^2 + rx^2)$ where the sign is chosen to make Ω a Sturm form. Take γ_i to be the subspace of $E \oplus E^*$ ($E =$ complex numbers) generated by $A^{\#}(b)[x_i]$. Then $x_i(t) = 0$ if and only if $\mathrm{nul}(\Omega, \infty \oplus \gamma_i, [t, b]) = 1$.

Since $\dim(E^\nu) = 1$ and $\dim(\gamma_1 \cap \gamma_2) \geqq 0$, the result follows.

8. Calculus of variations

Hermitian forms of the type $(\Omega, \beta, [a, b])$ occur in the calculus of variations as follows: Let M be a differentiable manifold. Let P be the space of all differentiable curves in M parametrized on $[0, 1]$. Let N be a submanifold of $M \times M$, and let $P_N = \{p \in P : (p(0), p(1)) \in N\}$. A function L on the tangent bundle $T(M)$ gives a function J on P by the usual formula $J(p) = \int_0^1 L(p'(t))dt$ where p' is the tangent vector to p. A path $p \in P_N$ will be called an *extremal of J relative to N* if the gradient of J is zero on the tangent space to P_N at p; or, more precisely, if $(d/ds)|_0 J[p_s] = 0$ whenever $s \to p_s \in P$ satisfies

(i) $(s, t) \to p_s(t)$ is differentiable for $(s, t) \in [-\varepsilon, \varepsilon] \times [0, 1]$.

(ii) $p_s \in P_N$ for $s \in [-\varepsilon, \varepsilon]$.

(iii) $p_0 = p$.

Now let p be a given extremal of J relative to N. Let B be the induced bundle $p^{-1}[T(M)]$; i.e., B is the real vector bundle on $[0, 1]$ whose fiber at $t \in (0, 1)$ is the tangent space to M at $p(t)$. Let $S(B)$ be the space of all differentiable cross-sections of B. [$S(B)$ is the tangent space to P at p and is best visualized as the space of vector fields along p.] Let $S(B)_N = \{x \in S(B) : (x(0), x(1))$ is tangent to N at $(p(0), p(1))\}$. [$S(B)_N$ is the tangent space to P_N at p.] Define a quadratic form H (the hessian) on $S(B)_N$ by

$$H[x] = \frac{d^2}{ds^2}\bigg|_{s=0} J[p_s]$$

where $s \to p_s$ satisfies (i)–(iii) above and has tangent vector x; i.e., $x(t)$ is the tangent vector to $s \to p_s(t)$ for each t.

PROPOSITION 8.1. *H is independent of the choice of p_s and defines a quadratic pair on $S(B)$ which is of the type $(\Omega, \beta, [0, 1])$ where order $\Omega \leqq 1$. [In particular, since domain $(\Omega, \beta, (0, 1]) =$ domain $H = S(B)_N$ we have that domain β is the tangent space to N at $(p(0), p(1))$].*

PROOF. This is the derivation of the "second variation formula," a straightforward computation consisting of differentiating under the integral sign and using the fact that p is an extremal of J relative to N.

The index of H is called the *index of p* considered as an extremal of J relative to N. The Morse theory relates the *indices of such extremals p* to the *topology of P_N*, subject to the assumption that L is such that the Ω of Proposition 8.1 is a Sturm form and, normally, the assumption that $\mathrm{nul}(H) = 0$ for all extremals p. Thus one is interested in *methods of computing indices of extremals*, which constitutes the second part of the Morse theory. Theorem

2.2 and Proposition 2.6 give $\mathrm{ix}(H)$ in terms of:

(1) The zeros of the solutions of Ω.

(2) The value of H on solutions of Ω.

(3) The quadratic form β.

or in more familiar terms:

(1) The zeros of Jacobi fields X along p.

(2) The values and covariant derivatives $(X(t), \dot{X}(t))$ of Jacobi fields at $t = 0$ and $t = 1$ (using formula (1.3) and the fact that $\dot{X}(t) = A(t)X$ in the usual formulation).

(3) The second fundamental form of $N \subset M \times M$ at $(p(0), p(1))$ relative to the normal direction given by p.

This method of compuation is valid subject to the following two remarks:

1. The Ω above is defined on a *real* vector bundle B. However, the problem is unchanged if it is complexified as is shown by the following:

PROPOSITION 8.2. *Let Q_R be a (real) quadratic form on a real vector space V_R, let $V_C = V_R \oplus i V_R$ be the complexification of V_R, and let Q_C be the unique extension of Q_R to a hermitian form on V_C (namely $Q_C[v + iw] = Q_R[v] + Q_R[w]$). Then $\mathrm{ix}(Q_C) = \mathrm{ix}(Q_R)$ and $\mathrm{nul}(Q_C) = \mathrm{nul}(Q_R)$.*

PROOF. The only part which is slightly tricky is the proof that $\mathrm{ix}(Q_C) \leqq \mathrm{ix}(Q_R)$. In the finite dimensional case, this is proved by writing Q_R in canonical form. The infinite dimensional case is reduced to the finite dimensional case by considering $W \vee \bar{W}$ for any finite dimensional subspace $W \subset V_C$ on which Q_C is negative definite.

2. Proposition 2.6 assumes that $\alpha \cap \infty = \{0\}$; i.e., no Jacobi fields vanish at both ends. For the general case we need a trivial improvement of this proposition:

PROPOSITION 8.3. *Let Ω be a Sturm form of order ν on $V[a, b]$. Let $\varepsilon \to \gamma(\varepsilon)$ for $\varepsilon \geqq 0$ be a $\oplus$-curve in $\mathrm{Herm}(E^\nu \oplus E^\nu)$ with $\gamma(0) = \alpha(\Omega, [a, b])$. Then for any $\beta \in \mathrm{Herm}(E^\nu \oplus E^\nu)$, we have*

$$\mathrm{ix}(\Omega, \beta, [a, b]) = \mathrm{ix}(\Omega, \infty, [a, b]) + \mathrm{ix}(\gamma(\varepsilon) - \beta \mid \textit{domain } \beta)$$

for all sufficiently small ε.

[For an application see Example I below].

PROOF. Let Λ be a Sturm form with $\int_a^b \Lambda$ positive definite and order $\Lambda \leqq$ order Ω. Set $\delta(\varepsilon) = \alpha(\Omega + \varepsilon\Lambda, [a, b])$. Then $\delta(\varepsilon)$ is a $\oplus$-curve (II of Theorem 3.1). By semi-continuity of the index (i.e., $\mathrm{ix}(\Omega + \varepsilon\Lambda, \beta, [a, b]) = \mathrm{ix}(\Omega, \beta, [a, b])$ for small ε) and by Proposition 2.6, the present proposition is true if we choose $\gamma(\varepsilon) = \delta(\varepsilon)$. But it is easily seen that the component of $\gamma(\varepsilon)$ in

$\mathrm{Herm}(E^\vee \oplus E^\vee) - \Gamma_\beta - \Gamma_\infty$ is independent of the choice of the $\oplus$-curve γ, hence so is $\mathrm{ix}(\gamma(\varepsilon) - \beta)$ by (IV) of Proposition 4.7, and the proposition follows.

We concluded with two examples considered by Bott [2]. The background of these is as follows:

Let M, L, J be given as before, and let N be the diagonal of $M \times M$. Then an extremal p of J relative to N is a closed curve and its n^{th} iterate g^n [i.e., go around g n-times] is easily seen (Euler's equations) to be an extremal of J relative to N too. The problem is to find the index of g^n. This is reduced to a set of boundary value problems by:

THEOREM (Bott). *Let Ω be a Sturm form of order* 1 *on* $V[0, n]$ *which is periodic with period* 1; i.e., $\Omega(t)[x] = \Omega(t-1)[y]$ *when* $x(t) = y(t-1)$. *For any complex number* $z \neq 0$, *define* $\beta_z \in \mathrm{Herm}(E \oplus E)$ *by domain* $B_z = \{(v, zv) : v \in E\}$ and $\beta_z[(v, zv)] = 0$. *Then*

$$\mathrm{nul}(\Omega, \beta_z, [0, n]) = \sum \mathrm{nul}(\Omega, \beta_\omega, [0, 1])$$
$$\mathrm{ix}(\Omega, \beta_z, [0, n]) = \sum \mathrm{ix}(\Omega, \beta_\omega, [0, 1]) \;.$$

where summation is over all n^{th} roots ω of z.

PROOF. By Theorem 3.1, it suffices to prove the statement for the nullity. This is done in [2, p. 177].

The problem of iterated extremals leads to the cases $z = \pm 1$, and hence to the functions Λ, N defined on the unit circle $|\omega| = 1$ by

$$\Lambda(\omega) = \mathrm{ix}(\Omega, \beta_\omega, [0, 1])$$
$$N(\omega) = \mathrm{nul}(\Omega, \beta_\omega, [0, 1]) \;.$$

[*Note*. A geometric problem leads to complex boundary conditions.] Now if $\alpha(\Omega, [0, 1])$ (henceforth denoted α) is given, then it is a matter of simple computation to find $N(\omega)$ and to find $\Lambda(\omega)$ up to an additive constant.

Computation of $N(\omega)$. As in Definition 2.7, let α be considered as a subspace of $F = E \oplus E^* \oplus E \oplus E^*$. (In the notation of Bott, α is the graph of the Poincaré matrix X.) Now $\beta_\omega \in \mathrm{Herm}(E \oplus E)$ is $\{(x, y, \omega x, \omega y\} \in F : x \in E, y \in E^*\}$ (using the fact that $\bar{\omega}^{-1} = \omega$) and $N(\omega) = \dim[\alpha \cap \beta_\omega]$ is easily found.

Computation of $\Lambda(\omega)$. Let $k = \mathrm{ix}(\Omega, \infty, [0, 1])$. Then $\Lambda(\omega) - k$ is given by Proposition 8.3. We consider the examples of [2, p. 181].

Example I. $\alpha = \beta_1$. In this case $N(\omega) = 0$ for $\omega \neq 1$ and $N(1) = 2n$ where $n = \dim E$. For $\gamma(\varepsilon)$ we can take the hermitian form on $E \oplus E$ defined by $\gamma(\varepsilon)[(x, y)] = \varepsilon \| x + y \|^2 - \varepsilon^{-1} \| x - y \|^2$, where $\| \; \|$ is some norm on E. Then $\Lambda(w) - k =$ index of $\gamma(\varepsilon)$ on $\{(x, \omega x)\}$ which is 0 if $\omega = 1$ and n if $\omega \neq 1$.

Example II. $\alpha[(x, y)] = (2\sigma)^{-1} \| y - x \|^2$, where $\| \; \|$ is a norm on E and $\sigma \neq 0$ is a real number. Then $N(\omega) =$ degree of degeneracy of α restricted to domain β_ω, which is 0 if $\omega \neq 1$, n if $\omega = 1$.

$$\Lambda(\omega) - k = \text{index of } \alpha \text{ on } \{(x, \omega x)\} = \begin{cases} 0 & \text{if } \omega = 1 \\ 0 & \text{if } \omega \neq 1,\ \sigma > 0 \\ n & \text{if } \omega \neq 1,\ \sigma < 0 \end{cases}$$

all of which agree with Bott's results.

COLUMBIA UNIVERSITY

BIBLIOGRAPHY

1. AMBROSE, W., *The index theorem in riemannian geometry*, Ann. of Math., 73 (1961), 49–86.
2. BOTT, R., *On the iteration of closed geodesics and the Sturm intersection theory*, Comm. Pure Appl. Math., 70 (1959), 313–337.
3. CODDINGTON, E. A. and LEVINSON, N., Theory of Ordinary Differential Equations, Mc-Graw-Hill, New York, 1955.
4. HEINZ, E., *Halbbeschränktheit Gewöhnlicher Differentialoperatoren Höherer Ordnung*, Math. Ann., 135 (1958), 1–49.
5. MORSE, M., The Calculus of Variations in the Large, AMS Colloquium Publications, v. 18, New York, 1934.
6. RAUCH, H. E., *A contribution to differential geometry in the large*, Ann. of Math., 54 (1951), No. 1, 38–55.

(Received March 11, 1963)