PACIFIC JOURNAL OF MATHEMATICS
Vol. 164, No. 2, 1994

CONJUGATE POINTS ON SPACELIKE GEODESICS OR PSEUDO-SELF-ADJOINT MORSE-STURM-LIOUVILLE SYSTEMS

ADAM D. HELFER

This paper develops the basic theory of conjugate points along geodesics in manifolds with indefinite metric; equivalently, that of conjugate points for Morse-Sturm-Liouville systems which are symmetric with respect to an indefinite inner product. The theory is rather different from that for Riemannian manifolds or that for timelike or null geodesics in Lorentzian manifolds. We find that conjugate points may be unstable with respect to perturbation of the geodesic: they may annihilate in pairs. Also the conjugate points need not be isolated: we construct an example where a whole ray is conjugate to a given point. Nevertheless, we give an extension of the Morse Index Theorem to this situation. We also analyze the effects of certain perturbations.

1. Introduction. The study of the length functional on Riemannian manifolds is fundamental to both classical and modern differential geometry. Classically, of course, the stationary points of this functional are the geodesics. The modern exploitation began with the Morse Index Theorem, which identified the index of the second variation (the "number of decreasing directions") with the algebraic count of the number of conjugate points along the geodesic. Morse himself used developments of this theory to prove deep results about the existence of periodic geodesics on the two-sphere equipped with an arbitrary metric [**11**]; Bott was led by a similar analysis of Lie groups to his celebrated Periodicity Theorem [**4**].

For Lorentzian manifolds, the existence of conjugate points on null or timelike geodesics has physical significance. For timelike geodesics, there is an effect rather like the "twin paradox," except that no accelerations are involved; for null geodesics, one has the phenomenon of gravitational lensing. Conjugate points along both these geodesics play a role in the singularity theorems of Penrose and Hawking, as well. (See [**13**] for a review.) The Morse Index Theorem and its consequences for these cases were established by Beem and Ehrlich [1–3]. We refer to the timelike and null cases collectively as *causal*.

Although the *interpretation* of the Morse Index Theorem is different in the Lorentzian causal cases from the Riemannian case and requires the development of new ideas [3], the result itself can be cast as a generalization of Sturm's theorem in such a form as to be identical in the Riemannian and causal Lorentzian cases.

MORSE INDEX THEOREM. *Consider the system of ordinary differential equations*

$$\frac{d^2}{dt^2}w^a = Q^a{}_b w^b \tag{1}$$

on the interval $[t_0, t_1]$, *where* $Q^a{}_b$ *is a matrix of smooth functions symmetric with respect to a positive-definite inner product* η_{ab}. *We say there is a conjugate point of multiplicity* k *at* $t = \tau$ *if there is a* k*-dimensional family of solutions vanishing at* $t = t_0$ *and* $t = \tau$. *Assume there is no conjugate point at* $t = t_1$. *Let*

$i_{\text{conjugate}}$, *the conjugate index, be the number of points in* (t_0, t_1) *conjugate to* t_0, *counted with multiplicity,*

i_{spectral}, *the spectral index, be the number of negative eigenvalues of the operator*

$$-\frac{d^2}{dt^2} + Q \tag{2}$$

on $[t_0, t_1]$ *with Dirichlet boundary conditions, counted with multiplicity,*

$i_{\text{quadratic}}$, *the index of the second variation, be the largest dimension of a subspace of the space of square-integrable vector-valued functions* $w(t)$ *on which the quadratic form*

$$\mathscr{Q}(w) = \int_{t_0}^{t_1} [\dot{w}^a \dot{w}^b \eta_{ab} + w^a w^b Q_{ab}]\, dt$$

(*where* $Q_{ab} = \eta_{ac} Q^c_b$) *is negative-definite.*

Then $i_{\text{conjugate}} = i_{\text{spectral}} = i_{\text{quadratic}}$.

(Sometimes the function space is taken to be something other than square-*integrable* functions; there is some latitude here.) Of course, one may choose a basis in which η_{ab} is the identity, and then Q_{ab} and $Q^a{}_b$ will have the same components. However, in what follows it will be important to maintain a conceptual distinction between the endomorphism $Q^a{}_b$ and the symmetric form Q_{ab}.

For the application to the length functional in a Riemannian manifold, the symbols have the following meaning. The interval $[t_0, t_1]$

parameterizes an interval along a geodesic γ; the vector-valued functions $w^a(t)$ take values in the normal bundle to the geodesic and are connecting vector fields to nearby geodesics; and $Q_{ab} = R_{arbs}\dot{\gamma}^r\dot{\gamma}^s$ is the Riemann curvature tensor contracted twice with the tangent to the geodesic. The symmetries of the Riemann tensor guarantee that $Q^a{}_b$ acts on the space of vectors normal to γ and is symmetric with respect to the induced metric on this normal bundle. At a point conjugate to $\gamma(t_0)$, the exponential map from $\gamma(t_0)$ fails to have maximal rank; in particular, the geodesic normal coordinates based at $\gamma(t_0)$ can only be a good chart on a neighborhood up to the first conjugate point. The eigenfunction corresponding to the most negative eigenvalue determines a canonical deformation of the geodesic which decreases its length most quickly. The index of the quadratic form is the index of the second variation: the number of independent perturbations decreasing the length.

For timelike geodesics in a Lorentzian manifold, essentially the same interpretations of the symbols apply. The only difference is that the metric η_{ab} on the normal bundle is now negative-definite (with our conventions a Lorentzian metric has diagonalized form $+1, -1, -1, \dots$), but the only modification needed to the Morse Index Theorem as stated above is to reverse the sign of $\mathcal{Q}$. Conjugate points still signal a drop in rank of the exponential map. The eigenfunction with the most negative eigenvalue now determines a canonical deformation of the geodesic *increasing* its length most quickly (the variational problem for timelike geodesics is to *maximize* their lengths; see [**3**, **12**]). The index of the quadratic form is the index of the second variation.

For null geodesics, the situation is a little more complicated. Here one must consider not the normal bundle, but the quotient of the normal bundle by the vectors proportional to $\dot{\gamma}^a$ (which, being null, is both tangent and normal). Once this is done, however, again η_{ab} is a negative-definite form and the Morse Index Theorem, as given above, can be applied. Conjugate points determine a drop in rank of the exponential map. The interpretation of the other two indices is a little more involved: see [**3**].

For spacelike geodesics, however, the form η_{ab} on the normal bundle is itself Lorentzian, and therefore the system of differential equations above is not self-adjoint with respect to a *definite* symmetric form. The Morse Index Theorem cannot be applied. This is tied to the fact that spacelike geodesics are *never* extrema of the length functional; indeed, there is always an infinite-dimensional family of deformations shortening the spacelike geodesic, and an infinite-dimensional family lengthening it: many of the usual techniques simply fail to apply here.

There has been little impetus from physics to investigate spacelike geodesics, since they are traversed by particles moving faster than light, which have never been observed and are believed not to exist. However, there is now, in addition to the mathematical motivation, a physical reason for studying these geodesics. This is that there are now good enough global existence theorems for Einstein's equation, due primilarly to Friedrich (see **[6]** for a review), that the focus of these investigations has shifted to spacelike infinity. This is a regime of ideal points which roughly speaking are endpoints of spacelike geodesics; it is now important to understand their geometry.

There is a further reason, both mathematical and physical, for studying the Morse Theory of spacelike curves. This is that it is the simplest of a family of variational problems encountered in relativity for which the functionals are bounded neither above nor below, and whose stationary points are never extrema. Virtually all of the relativistic field equations we have can be derived, and arise naturally, as stationary points of Lagrangians. At the moment, we are limited to rather formal uses of this fact, since there is not the mathematical technology available in this case analogous to what would exist for a minimization or maximization problem. One would like to know if it is possible to develop a theory complementary to that for positive-definite spaces. However, in general there may be two factors contributing to the unbounded nature of the relativistic variational problems, and only one is present in the treatment of spacelike geodesics. This is that the test functions take values in a space equipped with an indefinite metric. The other factor, which is not present in this paper, is that the differential equation characterizing stationarity may be hyperbolic, and hence the differential operator may not be semi-bounded.

In this paper, some of the basic theory of conjugate points on spacelike geodesics, and more generally on geodesics in semi-Riemannian spaces, is established. The first difference here with the causal Lorentzian or Riemannian theory is that *conjugate points on spacelike geodesics may be unstable.* That is, a perturbation of the system may destroy the conjugate points. Also *conjugate points may accumulate*; indeed, we give an example where a ray of points is conjugate to a fixed point. It turns out however that a version of the Morse Index Theorem does exist, and this is our main result.

We may view the geodesic deviation equation with appropriate initial conditions as giving rise, by a Legendre transformation, to a curve in the space of Lagrange planes. Then the conjugate index may be defined as the intersection number of this curve with a certain variety. In "generic" situations, this index is equal to the number of timelike conjugate points minus the number of spacelike conjugate points, where a conjugate point is classed as timelike or spacelike according to the value of $\dot{w}^a$ there. This conjugate index is equal to the spectral index, defined (generically) as the number of timelike minus the number of spacelike eigenfunctions of the operator

$$-\frac{d^2}{dt^2} + Q$$

(with Dirichlet boundary conditions) with negative eigenvalues, where an eigenfunction w^a is timelike or spacelike according to the sign of

$$\int_{t_0}^{t_1} w^a w_a \, dt.$$

The proof is by homology arguments in the space of Lagrange planes.

We give a number of related results analyzing the motion of conjugate points under certain perturbations. Inspecting these, it will be evident from the number of places signs enter that the theory is quite different from the definite case. We also show that our techniques extend to treat a somewhat more general class of problems, the Morse-Sturm systems.

This paper is entirely devoted to the study of the Morse Index Theory as a problem in ordinary differential equations. Applications of this to the geometry of space-time and Morse homology theory (as refined by Thom, Smale, Witten and others) will be given elsewhere.

This is the organization of the paper. In §2, we give the basic definitions that will be used. In §3, we show that every equation of the form (1) with $Q^a{}_b$ symmetric with respect to an indefinite form really is the geodesic deviation equation along some geodesic in a manifold, so all the phenomena discussed here can occur geometrically. Section 4 reviews the geometry of the Lagrange Grassmannian. Section 5 computes the conjugate index as an intersection number. Section 6 gives the spectral theory of the operator (2) and computes the spectral index. In §7, the index theorem is proved, and also an alternation theorem. Section 8 analyzes the effects of certain perturbations. Section 9 discusses focal points to hypersurfaces. Section 10 treats the Morse-Sturm problem. Section 11 gives some examples, in particular that of a point with a whole ray of conjugate points, and can be read before the other sections. Section 12 gives some final comments.

Acknowledgments. It is a pleasure to thank members of the Mathematics Department at the University of Missouri, Columbia, for useful discussions. I am particularly grateful to John Beem and to Fritz Gesztesy.

2. Preliminaries. The following symbols and terminology will be used.

The symbol $\odot$ stands for symmetric tensor product; the symbol $\cong$ for isomorphism of vector spaces; a dot will be used for d/dt.

A *null vector* will always mean a vector whose squared length, with respect to a given indefinite inner product, is zero. The term will not be used for an eigenvector with eigenvalue zero. A vector whose squared length is positive or negative will be called *timelike* or *spacelike* (with respect to the inner product).

The positive, negative and null *type numbers* of a quadratic form are the numbers of positive, negative and zero entries on the diagonal in its matrix of components with respect to a basis in which this matrix is diagonal. The *signature* of a quadratic form is its positive type number minus its negative type number.

The following definitions apply throughout this paper.

Let V be a real n-dimensional vector space. It will sometimes be convenient to denote the elements of the tensor algebra of V by quantities with indices in the usual fashion. When this is done, we represent the elements of V by symbols with small italic superscripts: v^a, w^a, etc.; and elements of the dual by quantities with subscripts: λ_a, etc. Contraction over repeated indices is understood. We will fix a nondegenerate symmetric bilinear form η on V. Indices will be raised and lowered with η and its inverse (also denoted η).

Let $Q_{ab}(t)$ be a smooth symmetric form. (For differential geometry, "smooth" may be taken to mean C^∞. However, for our purposes one could take Q_{ab} to be C^0.) We shall call the differential equation

$$\frac{d^2}{dt^2}w^a = Q^a{}_b w^b \tag{3}$$

a *Jacobi equation* symmetric with respect to η. A solution w^a to this equation is a *Jacobi field.* The *phase space* is defined to be $\Gamma = \{(w, \dot{w}) | w, \dot{w} \in V\} = V \oplus V$; it represents the specifiable data for a solution at any fixed t. The *symplectic form* on Γ is defined by

$$\omega((v, \dot{v}), (w, \dot{w})) = \eta(v, \dot{w}) - \eta(w, \dot{v}).$$

This form is preserved by evolution according to the differential equation: if $v^a(t)$ and $w^a(t)$ are solutions to (3), we have

$$\frac{d}{dt}\omega((v, \dot{v}), (w, \dot{w})) = 0$$

by virtue of the symmetry of Q_{ab}, independent of boundary conditions.

A generalization of the Jacobi equation (3) is the system

$$\frac{d}{dt}(r_{ab}(t)\dot{w}^b(t)) = Q_{ab}(t)w^b(t)$$

where $r_{ab}(t)$ is a C^1 non-degenerate symmetric form and Q_{ab} is C^0 and symmetric. We shall call this a *Morse-Sturm equation.* (So a Jacobi equation is the special case where $r_{ab} = \eta_{ab}$ is constant.) Although we are primarily interested in Jacobi equations, our techniques apply with only a few changes to Morse-Sturm equations, and a discussion of these is given below.

3. Every Jacobi equation is a geodesic deviation equation. It is clear that every geodesic deviation equation in a semi-Riemannian manifold is a Jacobi equation. We show that the converse is also true. Thus all the phenomena we describe for Jacobi equations can also occur for the geodesic derivation equation.

Let M be an $(n+1)$-dimensional manifold and g_{ab} a non-degenerate metric on M. Let O be an open interval of real numbers and $\gamma\colon O \to M$ an affinely parameterized geodesic. We recall that a Jacobi field on γ is a connecting vector field to a family of geodesics; it satisfies the *geodesic derivation equation*

$$(\dot{\gamma}\cdot\nabla)^2 w^d = R_{abc}{}^d\dot{\gamma}^a\dot{\gamma}^c w^b. \tag{4}$$

If we assume that γ is without self-intersections, then nothing is lost in pulling back the tangent bundle of M to γ, and trivializing this bundle by parallel propagation. We shall do this from now on. Then with $Q^d{}_b = R_{abc}{}^d\dot{\gamma}^a\dot{\gamma}^c$, the geodesic deviation equation becomes

$$\ddot{w}^d = Q^d{}_b w^b.$$

As a consequence of the symmetries of the Riemann tensor, any vector proportional to $\dot{\gamma}^a$ will be a solution of the geodesic deviation equation. Such solutions are not of interest for the theory of conjugate points. We therefore factor the space of solutions by these vectors. If $\dot{\gamma}^a$ is not a null vector, and we shall henceforth assume this, then we may equivalently require $w\cdot\dot{\gamma} = 0$. (This can always be arranged by adding a suitable multiple of $\dot{\gamma}$ to w.) We therefore consider (4) as an equation on the n-dimensional space of vectors w orthogonal to $\dot{\gamma}$; note that Q acts on this space anyway. Denote by η_{ab} the projection of the metric to the space orthogonal to $\dot{\gamma}$; then η is a non-degenerate symmetric form and Q is symmetric with respect to it.

We show now that every Jacobi equation does in fact arise in this manner. In other words, we show that every smooth function $Q^b{}_a$, symmetric with respect to η_{ab}, does arise as $R_{abc}{}^d\dot{\gamma}^a\dot{\gamma}^c$ for some semi-Riemannian manifold. This is a problem local to the geodesic: we must show that the metric can be chosen in the neighborhood of the geodesic so that its curvature has the requisite form.

This can be done as follows. Let (M, h_{ab}) be a semi-Euclidean space (that is, $M = \mathbf{R}^{n+1}$ and h_{ab} is a symmetric non-degenerate form, constant with respect to the Cartesian coordinate system). Let γ be a geodesic in (M, h_{ab}), timelike or spacelike as required, and let the type numbers of h_{ab} be such that the orthocomplement of $\dot{\gamma}$ has the same type numbers as η. We choose a conformally related metric $g_{ab} = \Omega^2 h_{ab}$, where

$$\Omega = 1, \quad \nabla_a\Omega = 0 \quad \text{on } \gamma.$$

(Here ∇_a is the covariant derivative with respect to h_{ab}.) Since $\nabla_a\Omega$ vanishes on γ, this curve will be a geodesic for g_{ab}. A direct calculation shows

$$Q_b{}^d = \text{the projection of } \dot{\gamma}^2\nabla_b\nabla^d\log\Omega \text{ orthogonal to } \dot{\gamma}^a$$

for g_{ab} on γ. We may choose this second derivative arbitrarily.

4. The Lagrange Grassmannian. The key step in the analysis of the Jacobi equation is to pass to the Lagrange Grassmannian, that is, the manifold of Lagrange planes in phase space. We review here those elements of the geometry of this space which will be needed below. Most of this material is standard, and for such results we give only as much of the proofs as will be necessary for understanding the remainder of this paper. For more details, see for example **[9, 14]**. We try to adhere to the conventions of these authors.

Recall that V is an n-dimensional real vector space equipped with a non-degenerate symmetric form η, and $\Gamma = V \oplus V$ is equipped with the symplectic form

$$\omega((v, \dot{v}), (w, \dot{w})) = \eta(v, \dot{w}) - \eta(w, \dot{v}).$$

An n-dimensional subspace L of Γ is called a *Lagrange plane* if ω restricts to zero on L. The space of Lagrange planes forms an $n(n+1)/2$-dimensional compact manifold Λ, called the *Lagrange Grassmannian.* Two Lagrange planes are *complementary* if their direct sum is Γ. We shall call two Lagrange planes which are not complementary *conjugate.* (Here and throughout we shall be careful to distinguish the capital lambda Λ, which is meant to suggest Lagrange, from the wedge $\wedge$, which denotes antisymmetry.)

If L_0, L_1 form a pair of complementary Lagrange planes, then the symplectic form can be used to identify each as the dual vector space of the other. We make the convention that the map

$$L_0 \to L_1^*$$

is given by

$$v \mapsto \omega(v\,,\,\cdot).$$

Although this is natural, it is not completely canonical; one might have chosen the map to be $v \mapsto \omega(\cdot\,,v) = -\omega(v\,,\cdot)$ instead. This will not be important. There is however a sign issue which is important: given our convention for the map $L_0 \to L_1^*$, note that the dual map $L_0^* \leftarrow L_1$ is given by $\omega(\cdot\,,v) \leftarrow\!\shortmid v$. Thus the identification $L_0 \to L_1^*$ really depends on the pair L_0, L_1: reversing their order introduces a sign change.

Similarly, for each Lagrange plane L we fix an isomorphism

$$\Gamma/L \to L^*$$

by

$$v + L \mapsto \omega(v\,,\,\cdot).$$

If L_0, L_1 form a complementary pair of Lagrange planes, this isomorphism and the one above are compatible in the sense that the diagram

$$\begin{array}{ccc} L_0 & \to & L_1^* \\ \downarrow & & \downarrow \\ \Gamma/L_1 & \to & L_1^* \end{array}$$

commutes, where the horizontal maps are the isomorphisms given above and the vertical maps are projection and the identity.

We now take up the structure of complementary pairs of Lagrange planes.

PROPOSITION 4.1. *Let L_0, L_1 be complementary Lagrange planes, and $e=(e_1, e_2, \ldots, e_n)$ a basis for L_0. Then there is a unique basis $f=(f^1, f^2, \ldots, f^n)$ for L_1 such that*

$$\omega(e_i, f^j)=\delta_i^j.$$

The basis f guaranteed here will be called *complementary* to e. In terms of components with respect to these bases, the canonical map $L_0 \to L_1^*$ is given by the identity. We have further

PROPOSITION 4.2. *Let L_0, L_1, e, f be as above. Then any Lagrange plane complementary to L_0 has a basis of the form*

$$f^i+\beta^{ij}e_j$$

for some unique symmetric matrix β^{ij}, and conversely any symmetric matrix determines a Lagrange complement to L_0 by this formula.

This shows that the set of Lagrange complements to L_0 is an affine space modeled on $L_0 \odot L_0$. If L_1 is chosen as the origin of the space, then the set of complements is identified with $L_0 \odot L_0$ by identifying a complement whose components are β^{ij} with $\beta^{ij}e_i \otimes e_j$. Let us define more generally, for any $L \in \Lambda$,

$$\Lambda^k(L)=\{M \in \Lambda \mid \dim L \cap M=k\}$$

for $k=0, 1, 2, \ldots, n$. Then $\Lambda^0(L)$ is the set of Lagrange complements to L, and $\Lambda=\bigcup_{k=0}^n \Lambda^k(L)$.

DEFINITION 4.1. Let L_0, L_1, e, f be as above. The *canonical chart* they determine is the diffeomorphism $\Lambda^0(L_0) \to \mathbf{R}^{n(n+1)/2}$ defined by $L \mapsto \beta^{ij}$ as above.

Now we identify the tangent vectors.

PROPOSITION 4.3. *The tangent space at $L \in \Lambda$ is canonically isomorphic to $L^* \odot L^*$.*

Proof. Take $L=L_1$, above. Then the tangent space is identified with the tangent space to an affine space modeled on $L_0 \odot L_0$. This tangent space may be identified with $L_0 \odot L_0$ itself, and we have L_0 canonically identified with L_1^*. (The isomorphism $T_L \cong L^* \odot L^*$ we have constructed does not depend on the sign of the identification of L_0 with L_1^*.) □

Note that with respect to the basis e for L_0, a tangent vector $\dot{L}$ at L_0 has components $\dot{L}_{ij}$. (So $\dot{L}=\dot{L}_{ij}e^{*i} \otimes e^{*j}$, where e^* is the dual basis to e.)

PROPOSITION 4.4. *Let* L_0, L_1 *be as above. Then the image of* $\Lambda^k(L_1)$ *in the canonical chart is the set of symmetric matrices of nullity* k. *This subspace, and hence* $\Lambda^k(L_1)$, *has codimension* $k(k+1)/2$ *in* Λ.

Proof. With e and f as above, let $L \in \Lambda^0(L_0)$ be given by $L = \operatorname{span}\{f^i + \beta^{ij} e_j\}$. Then $L_1 = \operatorname{span}\{f^i\}$, so

$$L \cap L_1 = \{\alpha_i f^i | \alpha_i \beta^{ij} = 0\}. \qquad \square$$

The codimension of $\Lambda^1(L)$ is one in Λ; and the codimension of $\Lambda^2(L)$ in $\Lambda^0(L)$ is three. We shall use these observations to construct the Arnol'd-Maslov cycle of L, which will be central to our later analysis. For any $L \in \Lambda$, let

$$A(L) = \Lambda - \Lambda^0(L).$$

We shall show that the variety $A(L)$ defines a cycle (in the sense of singular homology) and that this cycle has a natural transverse orientation. We noted above that in the sense of analytic varieties the regular set of $A(L)$ is $\Lambda^1(L)$ and the singular set has codimension two in $\Lambda^1(L)$ and codimension three in Λ. Thus $A(L)$ determines a cycle of codimension one in Λ. Since the codimension is one, in order to determine a transverse orientation it is enough to distinguish those transverse vectors which are positive from those which are negative. Let $M \in \Lambda^1(L)$, and suppose $\dot{M} \in T_M(\Lambda)$ is transverse to $\Lambda^1(L)$. This means that $\dot{M}$ viewed as a quadratic form on M does *not* restrict to zero on the one-dimensional subspace $L \cap M$. We define $\dot{M}$ to be positive if it restricts to a positive form on this space, and negative if it restricts to a negative form. This oriented cycle is the *Arnol'd-Maslov* cycle μ_L defined by L. The Arnol'd-Maslov cycle may be viewed as an element of $H^1(\Lambda, Z)$.

We now define the Maslov index of a curve. We need a slight generalization of the usual definition.

DEFINITION 4.2. Let $\gamma\colon [a, b] \to \Lambda$ be a curve with endpoints not on $A(L)$. Then the *Maslov index* of γ relative to L is the intersection number $\gamma \cdot \mu_L$.

If $\gamma\colon [a, b] \to \Lambda$ is a curve with $\gamma(b) \notin A(L)$ and $\gamma(a) \in A(L)$ but $\gamma(t) \notin A(L)$ for $t \in (a, a+\varepsilon]$ for some $\varepsilon > 0$, then the *Maslov index* of γ is defined to be the intersection number of $\gamma\colon [a+\varepsilon, b] \to \Lambda$ with $A(L)$.

(Of course, we could also consider curves which behave near $t = b$ in a fashion similar to that hypothesized in the second paragraph near

$t = a$, but we shall not need this.) If neither $\gamma(a)$ nor $\gamma(b)$ lie on $A(L)$, then the Maslov index is an invariant of homotopies of γ preserving $\gamma(a)$ and $\gamma(b)$. In the case of the second paragraph, the Maslov index is also a homotopy invariant for homotopies of curves satisfying the boundary conditions. That is, if $\gamma(t, s)\colon [a, b]\times[0, 1] \to \Gamma$ is continuous and for each $s \in [0, 1]$ we have

$$\begin{aligned} &\gamma(t, s) \neq A(L) \quad \text{for } t \in (a, a+\varepsilon_s), \\ &\gamma(a, s) = \gamma(a, 0), \qquad \gamma(b, s) = \gamma(b, 0), \end{aligned}$$

then the Maslov indices of $\gamma(\cdot, 0)$ and $\gamma(\cdot, 1)$ are equal. (The proof of this is elementary and will be omitted.)

If $\gamma\colon (a, b] \to \Lambda$ is smooth and intersects only the regular part of $A(L)$ and that transversely, the Maslov index is equal to the number of positive minus the number of negative tangents of γ at points of intersection. A more general result is the following.

PROPOSITION 4.5. *Let $\gamma\colon [a, b] \to \Lambda$ be a curve. Suppose that for some $\tau \in [a, b]$ we have $\gamma(\tau) \in A(L)$ and γ is one-to-one on an open interval containing τ. Choose L_0 complementary to both $\gamma(\tau)$ and L and let $L_1 = L$. Let $\varphi\colon \Lambda^0(L_0) \to L_0 \odot L_0$ be the canonical diffeomorphism defined by the pair L_0, L_1. Then the contribution to the Maslov index of γ at τ is one-half the change in signature of $\varphi\circ\gamma$ as t passes through τ.*

Proof. Let us write

$$\gamma(t) = \operatorname{span}\{f^i + \gamma^{ij}(t)e_j\}$$

as before. Suppose first that γ crosses $\Lambda^1(L_1)$ transversely at $t = \tau$. Then $\gamma^{ij}(\tau)$ has nullity one, but $\gamma^{ij}(t)$ has nullity zero for t near enough to, but unequal to τ. Let ξ_i be a non-zero column-vector in the kernel of $\gamma^{ij}(\tau)$, so

$$L \cap M = \operatorname{span}\{\xi_i, f^i\}.$$

Then by definition $\dot\gamma(\tau)$ is positive or negative according to whether $\dot\gamma^{ij}\xi_i\xi_j$ is. If $\dot\gamma$ is positive, then, the positive type number of γ^{ij} must increase by one as t increases from τ, and the negative type number decrease by one as t increases to τ. Thus if $\dot\gamma$ is positive, there will be a net change of $+2$ in the signature of γ^{ij} as t increases through τ. Similarly, if $\dot\gamma$ is negative there will be a net change of -2.

Now consider the case of a possibly non-transverse crossing. Let O be an open interval around τ small enough so that for $t \in O$

the curve intersects $A(L)$ only at $t = \tau$. Note that the signature of $\gamma^{ij}(t)$ is locally constant on $O - \{\tau\}$. Now perturb γ slightly in a closed subinterval F of O to a curve crossing only $\Lambda^1(L)$ and only transversely. At each crossing the signature will change by $+2$ or -2 according to whether the tangent is positive or negative. Thus, for the perturbed curve, the total contribution to the Maslov index will be one-half the change in signature. However, since the signature is locally constant on the complement of the image of $A(L)$, this change in signature for the perturbed curve must be the same as for the original.

5. The conjugate index. The Jacobi equation descends to Λ. We consider the system

$$\begin{cases} \ddot{w}_j^a = Q^a{}_b w_j^b, \\ w_j^a(t_0) = 0, \\ \dot{w}_j^a(t_0) = I_j^a, \end{cases}$$

where $I = (I_1, I_2, \dots, I_n)$ is a fixed basis for V. We regard this as a system of first-order equations for the pairs $(w_j, \dot{w}_j) \in \Gamma$. Since at $t = t_0$, the initial data determine a Lagrange plane, and since the symplectic form is preserved by evolution, we conclude that the solution to the system determines a curve $L(t)$ in Λ. This curve satisfies

$$L(t_0) = L_0 = 0 \oplus V.$$

If the space of w_j's satisfying $w_j(t) = 0$ has dimension k, we say t is conjugate to t_0 of multiplicity k. Evidently, this will be the same as $\dim L(t) \cap L_0 = k$, which is to say $L(t) \in \Lambda^k(L_0)$.

DEFINITION 5.1. Let $t \in [t_0, t_1]$, and suppose that the quadratic form $0 \oplus \eta$ restricted to $L(t) \cap L_0$ has positive, negative and null type numbers P, N and Z. We say there is a *timelike conjugate point* of multiplicity P, a *spacelike conjugate point* of multiplicity N, and a *null conjugate point* of multiplicity Z at parameter value t. The *signature* of the conjugate point is $P - N$.

The *multiplicity* of the conjugate point at t is $P+N+Z = \dim L(t) \cap L_0$. If the multiplicity is unity, we say there is a *simple* conjugate point at t. A simple conjugate point is said to be *timelike, spacelike* or *null* according as P, N or Z is unity.

Two remarks are in order. First, there is another notion of signature common in symplectic geometry, that of a triple of Lagrange planes. No confusion should arise. Second, *the multiplicity of the conjugate point at* t *may be strictly less than the order of vanishing of* $\det w_j^a$. An example will be given later.

Now let us work out the tangent to $L(t)$. We write

$$e_i(t) = (w_i(t), \dot{w}_i(t))$$

and let $f(t)$ be a complementary basis to $e(t)$. (Although we shall not need these formulae explicitly, we note that $f^i(t) = (u^i(t), \dot{u}^t(t))$ where $u^{ai}(t)$ is a solution of the equation $\ddot{u}^{ai} = Q^a{}_b u^{bi}$ with initial conditions determined by the requirement that $f(t_0)$ be a complementary basis to $e(t_0)$.) Then for τ near zero we may write

$$(w_i(t+\tau), \dot{w}_i(t+\tau)) = e_i(t) + L_{ij}(t+\tau) f^j(t)$$

with

$$\begin{aligned} L_{ij}(t+\tau) &= \omega(e_i(t+\tau), e_j(t)) \\ &= \eta(w_i(t+\tau), \dot{w}_j(t)) - \eta(\dot{w}_i(t+\tau), w_j(t)) \end{aligned}$$

the components of $L(t+\tau)$ with respect to the pair $(e(t), f(t))$. The components of the tangent vector with respect to this pair are

$$\begin{aligned} \frac{d}{d\tau} L_{ij}(t+\tau)\bigg|_{\tau=0} &= \eta(\dot{w}_i(t), \dot{w}_j(t)) - \eta(\ddot{w}_i(t), w_j(t)) \\ &= [w_i^a(t)\ \dot{w}_i^a(t)] \begin{bmatrix} -Q_{ab}(t) & 0 \\ 0 & \eta_{ab} \end{bmatrix} \begin{bmatrix} w_j^b(t) \\ \dot{w}_j^b(t) \end{bmatrix}. \end{aligned}$$

Thus $\dot{L}(t)$ is equal to the restriction of the bilinear form $-Q_{ab} \oplus \eta_{ab}$ on $V \oplus V$ to $L(t)$. We exploit this.

PROPOSITION 5.1. (a) *The curve* $L\colon [t_0, t_1] \to \Lambda$ *intersects* $\Lambda^1(L_0)$ *tranversely at* t *iff at this value of* t *there is a simple conjugate point; the intersection is positive or negative according to whether the conjugate point is timelike or spacelike.*

(b) *If every intersection of* $L(t)$ *with* $A(L_0)$ *is of finite order, then there are only finitely many conjugate points and the Maslov index of* $L(t)$ *is the sum of the signatures of the conjugate points.*

(c) *For an open dense set of* $Q^a{}_b(t)$*'s (in the* C^0 *topology), the conjugate points have only finite multiplicity; indeed* $\det w_j^a$ *has only simple zeros for generic* $Q^a{}_b$.

Proof. (a) We have

$$L(t) \cap L_0 = \{A^i(w_i^a, \dot{w}_i^a) | A^i w_i^a = 0\}.$$

The bilinear form $\dot{L}(t)$ restricts to $0 \oplus \eta_{ab}$ on elements of this space.

(b) Since $A(L_0)$ is compact, any infinite sequence of values of t for which $L(t) \in A(L_0)$ must have an accumulation point which must be an intersection of infinite order. Therefore under the present hypotheses there can be only finitely many conjugate points.

The second part of the claim follows from our discussion of the contributions to the Maslov index from general crossings of $A(L_0)$.

(c) By transversality, the zeros of a C^1 function on $[t_0, t_1]$ are generically simple. Since the function $\det w_j^a$ depends continuously on the function $Q^a{}_b$ (holding fixed the initial conditions), the set of $Q^a{}_b$'s for which the zeros of $\det w_j^a$ are simple is open. We now argue that this set must be dense as well.

Let $Q^a{}_b$ be given. We may perturb it by an arbitrarily small amount to an analytic function. Then the perturbed w_j^a and hence the perturbed $\det w_j^a$ will be analytic. Then $\det w_j^a$ can have only finitely many zeros on $[t_0, t_1]$, and those which it does have can only have finite order. We now show that by a further arbitrarily small analytic perturbation, we can make each of the zeros (which survives the perturbation) simple. It is enough to show that there is a perturbation destroying the degeneracy of any given zero (since then a finite sequence of such perturbations can be used to destroy all multiplicities). However, if this were not true, then there would be one zero whose degeneracy was preserved by arbitrary (sufficiently small) analytic perturbations. However, by analyticity, the degeneracy of this zero would then be preserved by arbitrary analytic perturbations. This is clearly impossible. □

COROLLARY. *At a simple null conjugate point, the curve L cannot be transverse to $\Lambda^1(L_0)$, and $\det w_j^a$ must vanish to order greater than one.*

Motivated by these results, we make the

DEFINITION 5.2. Assume there is no conjugate point at $t = t_1$. The *conjugate index* of the Jacobi equation, denoted $i_{\text{conjugate}}$, is the Maslov index of L.

For the initial conditions we are considering, we are guaranteed there is no conjugate point in some interval $(t_0, t_0 + \varepsilon]$, so the conjugate index is well-defined.

6. Spectral theory and the spectral index. We begin with a result in the finite-dimensional case, which is of some interest in the present situation.

PROPOSITION 6.1. *Let $S^a{}_b$ be an endomorphism on V, symmetric with respect to η. Then there is a (complex) basis for V in which $S^a{}_b$*

takes the Jordan canonical form and the metric is block-diagonal with blocks of the same size and location as the Jordan blocks, each block of the metric being of the form

$$\pm \begin{bmatrix} & & & & 1 \\ & & & 1 & \\ & & \cdot^{\cdot^{\cdot}} & & \\ & 1 & & & \\ 1 & & & & \end{bmatrix}$$

(*blank places are occupied by zeros*).

See for example [7].

Two points are worth mentioning. First, the possible sizes of the blocks and combinations of $\pm$ signs are restricted by the signature of η. For example, in the case of physical interest for this paper, spacelike geodesics in a four-dimensional Lorentzian space-time, the only possibilities for η are

$$\begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix}, \begin{bmatrix} -1 & & \\ & & 1 \\ & 1 & \end{bmatrix}, \begin{bmatrix} -1 & & \\ & & -1 \\ & -1 & \end{bmatrix}, \begin{bmatrix} & & -1 \\ & -1 & \\ -1 & & \end{bmatrix}$$

(and permutations of the blocks in these). Second, the direct sum decomposition defined by the Jordan blocks is an orthogonal direct sum.

We now turn to the eigenvalue problem for the Jacobi equation with Dirichlet boundary conditions. There is some choice in the function space to be used for the operator-theoretic analysis. The freedom involved is not significant for us, and it will be technically simplest to use an analog of the familiar Hilbert-space analysis. Fix an arbitrary positive-definite form h_{ab} (independent of t), and consider the space of V-valued functions which are in the L^2 space defined by h_{ab} on $[t_0, t_1]$. This topological vector space, which we denote by H, is independent of the choice of h_{ab}. It is therefore a *Hilbertable space*, that is, a topological vector space which is topologically isomorphic to Hilbert space, but not equipped with any preferred norm. In fact, we equip H with the indefinite norm

$$\|w\|^2 = \int_{t_0}^{t_1} \eta_{ab} w^a w^b \, dt.$$

We write the inner product on H as

$$\langle u , w\rangle = \int_{t_0}^{t_1} \eta_{ab} u^a w^b \, dt.$$

Then the inner product is a continuous bilinear form, and the map $w \mapsto \langle w , \cdot\rangle$ determines a topological isomorphism $H \to H^*$. An element of H is said to be *timelike, null* or *spacelike* according to whether its squared norm is positive, zero or negative.

To treat the eigenvalue problem

$$\begin{cases} -\ddot{w}^a + Q^a{}_b w^b = \lambda w^a , \\ w^a(t_0) = 0 , \\ w^a(t_1) = 0 , \end{cases}$$

it will be convenient to consider the differential equation

$$\begin{cases} -\ddot{w}^a_j + Q^a{}_b w^b_j = \lambda w^a_j , \\ w^a_j(t_0) = 0 , \\ \dot{w}^a_j(t_0) = I^a_j . \end{cases} \tag{5}$$

(The function w^a_j is thus understood to depend on λ, although we do not usually write this dependence explicitly.) Then λ_0 is an eigenvalue with eigenspace of dimension k iff $w^a_j(t_1)$ has nullity k at $\lambda = \lambda_0$. Thus no eigenspace has dimension more than n.

We also note the following elementary facts. First, $w^a_j(t_1)$ and $\det w^a_j(t_1)$ are analytic functions of λ; hence they have zeros of finite multiplicities only, and so eigenvalues can accumulate only to infinity. Second, because $Q^a{}_b$ is bounded as a linear operator on the Hilbertable space, the real parts of the eigenvalues are bounded below by $\pi^2/(t_1 - t_0)^2 - \|Q\|$ (here $\|Q\|$ is the operator norm with respect to any Hilbert structure) and the imaginary parts of the eigenvalues are bounded in magnitude by $\|Q\|$. (It is quite possible that there are complex eigenvalues.) Third, the dimension of an eigenspace may be srtictly less than the order of the zero of $\det w^a_j(t_1)$ as a function of λ.

There is a spectral theorem for this situation.

THEOREM 6.1. *Let J be the operator $-d^2/dt^2 + Q$ on $D = \{w \in H \mid w$ is C^2 and $w(t_0) = w(t_1) = 0\}$. Then the eigenvalues of J are isolated, have their imaginary parts bounded and their real parts bounded below. There is an associated resolution of the identity*

$$1 = \sum_{\text{eigenvalues}} E_\lambda$$

where each E_λ is projection onto a finite-dimensional subspace H_λ of H (or its complexification if λ is complex). If $\lambda \neq \mu$ then H_λ and H_μ are orthogonal.

Let $H_J = \{w \in H | w$ lies in a direct sum of finitely many H_λ 's$\}$. If $w \in H_J$ we have

$$Jv = \left[\frac{1}{2\pi i} \oint \frac{\lambda}{\lambda - J} d\lambda\right] v$$

for a contour enclosing sufficiently many eigenvalues, each once in the positive sense. In particular, on H_J the operator J is equal to $\sum_\lambda J_\lambda$ where

$$J_\lambda = \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi i} \oint_{|\xi - \lambda| = \varepsilon} \frac{\xi}{\xi - J} d\xi.$$

Proof. That J possesses a resolution of the identity is standard [5]. To see that each projector has finite-dimensional range, imagine writing down the resolvent kernel from solutions of the homogeneous problem by variation of parameters in the usual way. The only obstruction to doing so (which signals the presence of an eigenvalue) is the need to solve some algebraic "matching equations." The terms in these equations depend analytically on λ, and so the resolvent can have at most a pole of finite order at any value of λ.

Now suppose $w \in H_\lambda$ and $v \in H_\mu$. Then there are positive integers n, m so that $(J - \lambda)^n w = 0$ and $(J - \mu)^m v = 0$. For any $N \geq n$ we have

$$\begin{aligned} 0 &= \langle (J - \lambda)^N v , w \rangle = \langle v , (J - \lambda)^N w \rangle \\ &= \langle v , (J - \mu + \mu - \lambda)^N w \rangle \\ &= \sum_{k=0}^{n-1} \binom{N}{k} (\mu - \lambda)^{N-k} \langle v , (J - \mu)^k w \rangle. \end{aligned}$$

It is now an exercise to show that this implies $\langle v , (J - \mu)^k w \rangle = 0$ for $k = 0, 1, \dots, n - 1$.

The remaining claims follow from standard analysis [5]. □

The operator J extends naturally to the Sobolev space of V-valued functions vanishing at t_0 and t_1 and whose second derivatives are h_{ab}-square-integrable on $[t_0, t_1]$; for w in this space the sum $\sum_\lambda J_\lambda w$ is unconditionally convergent. In particular, we may study the space of w 's which are C^0 and piecewise C^2, which is common in the classical treatment of the Morse Index Theorem. (If one is willing to extend the range of J beyond L^2, one can extend J to the Sobolev space of V-valued functions vanishing at t_0 and t_1 whose first derivatives are h_{ab}-square-integrable.) We remark that it is quite possible that J has a quasi-nilpotent part, so that some of the $J_\lambda - \lambda$'s are nilpotent. The structure of J_λ is determined by applying the proposition above.

We are now in a position to define the spectral index for the Jacobi problem.

DEFINITION 6.1. Let λ be an eigenvalue of the Jacobi equation, and H_λ the corresponding subspace in H. We call the positive, null and negative type numbers of $\langle \cdot , \cdot \rangle$ restricted to H_λ the *timelike, null* and *spacelike* multiplicities of λ.

The *spectral index* i_{spectral} of the Jacobi problem is the number of timelike negative eigenvalues, counted with multiplicity, minus the number of spacelike negative eigenvalues, counted with multiplicity.

PROPOSITION 6.2. *For each real λ, let $L(t, \lambda)$ be the curve in Λ corresponding to the differential equation* (5). *Assume that there is no conjugate point at $t = t_1$ (for $\lambda = 0$). Then i_{spectral} is equal to the Maslov index of the curve $L(t_1, \cdot)$ as λ varies from a sufficiently negative λ_- to zero.*

Proof. The argument is similar to that of the previous section. We must compute the intersection number of the curve $L(t_1, \lambda)$ with $A(L_0)$. In the following calculation the index i is not summed over.

$$\begin{aligned}
\omega(\partial_\lambda e_i , e_i) &= \eta(\partial_\lambda w_i(t_1), \dot{w}_1) - \eta(\partial_\lambda \dot{w}_i(t_1), w_i(t_1)) \\
&= \int_{t_0}^{t_1} [\ddot{w}_i^a \partial_\lambda w_{ai} - w_i^a \partial_\lambda \ddot{w}_{ai}]\, dt \\
&= \int_{t_0}^{t_1} [\ddot{w}_i^a \partial_\lambda w_{ai} - w_i^a (Q_{ab} \partial_\lambda w_i^b - \lambda \partial_\lambda w_{ai} - w_{ai})]\, dt \\
&= \int_{t_0}^{t_1} [\partial_\lambda w_{ai} (\ddot{w}_i^a - Q^a{}_b w_i^b + \lambda w_i^a) + w_{ai} w_i^a]\, dt \\
&= \|w_i\|^2.
\end{aligned}$$

By polarization, then, the quadratic form $\partial_\lambda L(t_1, \lambda)$ is simply $\langle \cdot , \cdot \rangle$.

7. The Index Theorem. We can now prove the extension of the Morse Index Theorem.

THEOREM 7.1. *The conjugate and spectral indices for the Jacobi problem are equal. Also there exists an orthogonal direct sum decomposition $H = T \oplus S$ into timelike and spacelike subspaces for which the* *restriction of the second variation to T has finite negative type number, and the restriction to S has finite positive type number, and the difference in these two type numbers is again equal to the index.*

Proof. For each real λ, let $L(t, \lambda)$ be the curve in Λ defined by the Jacobi equation. For λ negative enough we are guaranteed no conjugate points in $(t_0, t_1]$. Suppose λ_- a real number such that there are no conjugate points in $(t_0, t_1]$ for $\lambda \le \lambda_-$. Also choose $\varepsilon > 0$ small enough so that there is no conjugate point for $t \in (t_0, t_0 + \varepsilon]$ for any $\lambda \in [\lambda_-, 0]$. Then we have a simplex

$$\{(t, \lambda) \in [t_0 + \varepsilon, t_1] \times [\lambda_-, 0]\} \to \Lambda$$

given by

$$(t, \lambda) \mapsto L(t, \lambda).$$

Two of the boundary curves,

$$L(\cdot, 0)\colon [t_0 + \varepsilon, t_1] \to \Lambda \quad \text{and} \quad L(t_1, \cdot)\colon [\lambda_-, 0] \to \Lambda,$$

have intersection numbers with $A(L_0)$ we want to compare. The remaining two boundary curves,

$$L(t_0 + \varepsilon, \cdot)\colon [\lambda_-, 0] \to \Lambda \quad \text{and} \quad L(\cdot, \lambda_-)\colon [t_0 + \varepsilon, t_1] \to \Lambda$$

have zero intersection number with $A(L_0)$, by construction. □

The final statement here is an attempt to link the indices given so far with the variational problem. The development of this theory will be given elsewhere.

One of the results of the classical Sturm-Liouville theory is that zeros of the solutions alternate with zeros of their derivatives. Here is an analog.

THEOREM 7.2. *Let w_j^a be an n-dimensional family of Jacobi fields on $[t_0, t_1]$, and suppose that at neither end-point is the associated Lagrange plane conjugate to either $L_0 = 0 \oplus V$ or $L_1 = V \oplus 0$. Let $L(t)$ be the curve of Lagrange places associated to the Jacobi fields. Then*

$$|L \cdot \mu_{L_0} - L \cdot \mu_{L_1}| \le n.$$

Proof. First note that as $M \in \Lambda$ varies continuously, so does the cycle $A(M)$. Since Λ is connected, then, any two Arnol′d-Maslov cycles are homologous.

Let γ be a path in Λ from $L(t_1)$ to $L(t_0)$. (We shall specify γ more precisely soon.) Then $L+\gamma$ is a 1-cycle. By our observation of the previous paragraph, $(L+\gamma)\cdot(\mu_{L_0}-\mu_{L_1})=0$. Thus

$$L\cdot\mu_{L_0}-L\cdot\mu_{L_1}=\gamma\cdot\mu_{L_1}-\gamma\cdot\mu_{L_0}.$$

We show that we can choose γ so that the right-hand side has absolute value no larger than n.

By hypothesis, the end-points of γ lie in $(\Lambda-A(L_0))\cap(\Lambda-A(L_1))=\Lambda^0(L_0)\cap\Lambda^0(L_1)$. In a canonical chart the image of $\Lambda^0(L_0)$ is the space of symmetric $n\times n$ matrices and the image of $\Lambda^0(L_0)\cap\Lambda^0(L_1)$ is the space of non-singular $n\times n$ symmetric matrices. Any two such non-singular matrices can be joined by a path over which half the net change in signature is no more than $\pm n$.

8. Effects of certain perturbations. We consider here the motion of the conjugate points when the Jacobi equation is perturbed in certain ways. It is possible to strengthen these results somewhat, but because of the possible degeneracies the formulations grow a bit involved, and we give only the simplest, cleanest, results.

Some elementary observations first: If there is no conjugate point at $t=t_1$, then under small enough perturbations of $Q^a{}_b$, the conjugate index is stable. Also, for small enough perturbations, at an isolated conjugate point, the order of contact of $L(t)$ with $A(L_0)$ cannot increase if it is finite. (This from general transversality theory.)

Here is a comparison result.

THEOREM 8.1. *Let Q_{ab} be perturbed by the addition of a function which is positive-definite for all t. Then the isolated timelike conjugate points move left and the isolated spacelike conjugate points move right.*

Proof. We consider the Jacobi equation

$$\ddot{w}^a=(Q^a{}_b+sh^a{}_b)w^b$$

where $h_{ab}(t)$ is positive-definite with the same initial conditions as before. Then $\partial_s w^a_j$ satisfies

$$\begin{cases}\partial_s\ddot{w}^a_j=Q^a{}_b\partial_s w^b_j+h^a{}_b w^b_j\,,\\ \partial_s w^a_j|_{t=t_0}=0\,,\\ \partial_s\dot{w}^a_j|_{t=t_0}=0.\end{cases}$$

Let $L(t, s)$ be the corresponding curve in Λ. Then

$$\begin{aligned}\partial_s L_{ij} &= \eta(\partial_s w_i, \dot{w}_j) - \eta(\partial_s \dot{w}_i, w_j) \\ &= -\int_{t_0}^{t} h_{ab} w_i^a w_j^b \, dt\end{aligned}$$

so $\partial_s L$ is a negative-definite quadratic form. Now consider the change in L at an isolated conjugate point as s is increased from zero. Say we choose L_1 complementary to L_0 and the image of the conjugate point in Λ as usual, and identify as usual $A(L)$ as the space of degenerate elements in $L_0 \oplus L_0$. Indeed, choose a specific basis for L_0, and compute the eigenvalues of the matrix L_{ij}. As, for $s = 0$, the parameter t passes through the conjugate value, some of these eigenvalues will change sign. However, for positive values of s, the sign changes from negative to positive will occur sooner, and those from positive to negative will occur later. □

Similarly, if Q_{ab} is perturbed by the addition of a negative-definite form, then isolated timelike conjugated points move right and isolated spacelike conjugate points move right.

It is also of interest to know how the conjugate points vary with the initial point.

THEOREM 8.2. *Suppose there is a simple conjugate point at $t = \tau$. Then as t_0 moves to the right, this point moves to the right or the left according to whether $\dot{w}^a(t_0)\dot{w}_a(t_0)$ and $\dot{w}^a(\tau)\dot{w}_a(\tau)$ have the same or opposite signs, where w^a is the nontrivial Jacobi field vanishing at $t = \tau$. (If $\dot{w}^a(t_0)\dot{w}_a(t_0) = 0$, the conjugate point is stationary to first order.)*

Proof. Suppose for the moment that Q is of class C^1. It is technically easier to move Q to the left: let $Q^a{}_b(t, s) = Q^a{}_b(t+s)$ for s positive. Let $L(t, s)$ be the corresponding curve in Λ. If the conjugate point occurs at $t = \tau(s)$, then we wish to find whether $s + \tau(s)$ moves to the right or the left as s increases, that is, to find the sign of $1 + \partial_s \tau$. Since we have $L(\tau(s), s) \in \Lambda^1(L_0)$, we find

$$\partial_t L - \partial_s L = (1 + \partial_s \tau)\partial_t L + \text{terms tangent to } \Lambda^1(L_0).$$

Thus we may find the sign of

$$\lambda^{ij}(\partial_t L_{ij} - \partial_s L_{ij})$$

relative to $\lambda^{ij}\partial_t L_{ij}$, where λ^{ij} are the components of the covector normal to $\Lambda^1(L_0)$.

We have already computed

$$\partial_t L = \begin{bmatrix} -Q_{ab} & \\ & \eta_{ab} \end{bmatrix}$$

as a bilinear form on L_0, and λ at the conjugate point is evaluation of this form on

$$\begin{bmatrix} 0 & \\ & \dot{w}^a(\tau)\dot{w}^b(\tau) \end{bmatrix}$$

where w is as specified in the hypothesis. Thus

$$\lambda^{ij}\partial_t L_{ij} = \eta_{ab}\dot{w}^a(\tau)\dot{w}^b(\tau).$$

From the calculation of the previous theorem, we find

$$\partial_s L_{ij} = -\int_{t_0}^{t} \dot{Q}_{ab} w_i^a w_j^b \, dt.$$

Thus

$$\begin{aligned} \partial_t L_{ij} - \partial_s L_{ij} &= -Q_{ab} w_i^a w_j^b + \dot{w}_i^a \dot{w}_{aj} + \int_{t_0}^{t} \dot{Q}_{ab} w_i^a w_j^b \, dt \\ &= \dot{w}_i^a \dot{w}_{aj} - \int_{t_0}^{t} Q_{ab}(\dot{w}_i^a w_j^b + w_i^a \dot{w}_j^b)\, dt \\ &= \dot{w}_i^a \dot{w}_{aj} - \int_{t_0}^{t} (\dot{w}_i^a \ddot{w}_{aj} + \ddot{w}_i^a \dot{w}_{aj})\, dt \\ &= \dot{w}_i^a \dot{w}_{aj} - \int_{t_0}^{t} \frac{d}{dt}(\dot{w}_i^a \dot{w}_{aj}) = \dot{w}_i^a(t_0)\dot{w}_{aj}(t_0). \end{aligned}$$

Thus $\lambda^{ij}(\partial_t L_{ij} - \partial_s L_{ij}) = \dot{w}^a(t_0)\dot{w}_a(t_0)$.

Finally, if Q is only assumed to be C^0, this computation is still valid if $\dot{Q}_{ab}$ is interpreted distributionally. □

If t_0 moves to the left, of course, the behavior of the conjugate points is the opposite.

9. Focal points to hypersurfaces. Our techniques can be used to treat the Jacobi equation with other initial conditions. We discuss here the case of most geometric interest, focal points to a hypersurface along an orthogonal geodesic. (These are also sometimes called conjugate points to the hypersurface.) The Jacobi equation is

$$\begin{cases} \ddot{w}_j^a = Q^a{}_b w_j^a, \\ w_j^a(t_0) = I_j^a, \\ \dot{w}_j^a(t_0) = 0, \end{cases}$$

with I_j^a as before. Thus the curve $L(t)$ has as its origin

$$L(t_0) = V \oplus 0;$$

however, we are still interested in Lagrange planes conjugate to $L_0 = 0 \oplus V$; these occur when there are *focal points* of hypersurface along the geodesic. In this case the focal index (defined in the same way as the conjugate index) is equal to the spectral index for the mixed boundary problem

$$\begin{cases} -\ddot{w}^a + Q^a{}_b w^b = \lambda w^a, \\ \dot{w}^a(t_0) = 0, \\ w^a(t_1) = 0. \end{cases}$$

We find again that the addition of a positive-definite perturbation to Q_{ab} causes timelike conjugate points to move left and spacelike conjugate points to move right. The calculation is the same as previously; it is only necessary to check that no new boundary terms are introduced. On the other hand, the behavior of conjugate points as t_0 is moved is different.

THEOREM 9.1. *Suppose there is a simple focal point at $t = \tau$. Then as t_0 moves to the right, this point moves to the right or the left according to whether $-Q_{ab}(t_0)w^a(t_0)w^b(t_0)$ and $\dot{w}^a(\tau)\dot{w}_a(\tau)$ have the same or opposite signs. (Here w is the non-trivial Jacobi field with focal point at $t = \tau$.)*

Proof. This is the same as before, but with the present boundary conditions we find

$$\partial_t L_{ij} - \partial_s L_{ij} = -Q_{ab}(t_0)w_i^a(t_0)w_j^b(t_0).$$

10. The Morse-Sturm problem. In this section we indicate the changes necessary to treat the Morse-Sturm equation

$$\frac{d}{dt}(r_{ab}(t)\dot{w}^b) = Q_{ab}w^b$$

on the interval $[t_0, t_1]$, where r_{ab} is a smooth non-degenerate symmetric form.

The phase space is here $\Gamma = V \oplus V^*$, with the data for the Morse-Sturm equation at time t defining a point $(w^a(t), r_{ab}(t)\dot{w}^b(t))$ in Γ. The preserved symplectic form is

$$\omega((v^a, r_{ab}\dot{w}^b), (w^a, r_{ab}\dot{w}^b)) = v^a r_{ab}\dot{w}^b - w^a r_{ab}\dot{w}^b.$$

The signature of a conjugate point at parameter value $t = \tau$ must now be defined as the signature of the form $0 \oplus r_{ab}(\tau)$ restricted to $L(\tau) \cap L_0$.

The norm on the Hilbertable space must now be defined as

$$\|w\|^2 = \int_{t_0}^{t_1} r_{ab}(t) w^a w^b \, dt$$

with the inner product determined by polarization. The spectral theory is the same except for the obvious change in the domain of the operator and the equality of the conjugate and spectral indices is established as before. So too for the motion of conjugate points under perturbations. Additionally, we have the following.

THEOREM 10.1. *Let r_{ab} be perturbed by the addition of a positive-definite function (but remain non-degenerate). Then isolated timelike conjugate points move left and isolated spacelike conjugate points move right.*

Proof. The argument is analogous to that for the perturbation of Q_{ab}. If we replace r_{ab} by $r_{ab} + s\rho_{ab}$, then we find

$$\partial_s L_{ij} = -\int_{t_0}^{t_1} \rho_{ab} \dot{w}_i^a \dot{w}_j^b \, dt.$$

11. Examples. The examples we give are all in the case $n = 2$ with η_{ab} Lorentzian. We shall begin with the case where Q_{ab} is constant.

A *null basis* is one in which the metric takes the form

$$\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}.$$

From the general classification of symmetric forms, we conclude that one of the following holds:

(a) There is a null basis in which Q_{ab} has the form

$$\begin{bmatrix} 0 & \lambda \\ \lambda & \kappa \end{bmatrix}.$$

(b) There is an orthonormal basis in which Q_{ab} has the form

$$\begin{bmatrix} \mu & \nu \\ \nu & \mu \end{bmatrix}.$$

(c) There is an orthonormal basis in which the matrix is diagonal.

In alternative (a), the eigenvalue λ occurs with multiplicity two; in (b), there are complex eigenvalues $\mu \pm i\nu$.

Now let us consider the system

$$\ddot{w}^a = Q^a{}_b w^b$$

for $Q^a{}_b$ a constant matrix. Then the solutions satisfying the initial conditions $w(0) = 0$, $\dot{w}(0) = I$ (where I is the identity) are: for case (a)

$$w = \begin{bmatrix} \frac{1}{\sqrt{\lambda}} \sin \sqrt{\lambda} t & \frac{\kappa}{2\lambda} t \cos \sqrt{\lambda} t \\ 0 & \frac{1}{\sqrt{\lambda}} \sin \sqrt{\lambda} t \end{bmatrix}$$

for case (c),

$$w = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} \sin \sqrt{\lambda_1} t & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \sin \sqrt{\lambda_2} t \end{bmatrix}$$

and for case (b), a formula like that for case (c) in a complex orthonormal basis, with λ_1 and λ_2 given by λ and $\overline{\lambda}$.

There will be points conjugate to $t = 0$ if and only if at least one eigenvalue is real and positive. This is not a stable condition. By an arbitrarily small perturbation of case (a) (or of case (c) with $\lambda_1 = \lambda_2$), we can arrive at case (b), that of complex eigenvalues. Notice however that under such a perturbation the conjugate points annihilate in pairs.

We now show that *conjugate points may accumulate.* Indeed, we give an example in which every $t \geq 1$ is conjugate to $t = 0$.

In order to construct this, it will be helpful to work backwards from the equation

$$\ddot{w} = Qw.$$

We shall give w, and then verify that Q defined by $Q = \ddot{w}w^{-1}$ where w is nonsingular and by a limit of this elsewhere is smooth and symmetric. Let us put

$$w = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

in a null basis. Then the form

$$\begin{aligned} Q^{ab} &= \ddot{w}^a_i (w^{-1})^i_r \eta^{rb} \\ &= \begin{bmatrix} \ddot{\alpha} & \ddot{\beta} \\ \ddot{\gamma} & \ddot{\delta} \end{bmatrix} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} \end{aligned}$$

must be symmetric. This implies

$$\ddot{\alpha}\delta - \ddot{\beta}\gamma = -\beta\ddot{\gamma} + \alpha\ddot{\delta},$$

that is,

$$\frac{d}{dt}(\dot{\alpha}\delta - \alpha\dot{\delta}) = \frac{d}{dt}(\dot{\beta}\gamma - \beta\dot{\gamma}).$$

At $t = 0$, both quantities in parentheses are zero, so

$$\dot{\alpha}\delta - \alpha\dot{\delta} = \dot{\beta}\gamma - \beta\dot{\gamma}$$

which we regard as a differential equation for δ in terms of given functions α, β, γ. (These are not quite arbitrarily specifiable, since Q is required to be smooth.) With our initial conditions, the solution to this equation is

$$\delta = \alpha \left[\int_0^t \alpha^{-2}(\beta\dot{\gamma} - \gamma\dot{\beta})\, dt + 1 \right].$$

Now choose α, β and γ as follows:

(a) On $[0, 1/4]$, let $\alpha = t$; on $[1/2, \infty)$, let $\alpha = 1$, and on $(1/4, 1/2)$ interpolate smoothly and monotonically.

(b) On $[0, 1/2]$, let $\beta = 0$; on $[3/4, \infty)$, let $\beta = t - 1$, and on $(1/2, 3/4)$ interpolate smoothly by a negative function.

(c) On $[0, 3/4]$ and on $[1, \infty)$, let $\gamma = 0$; on $(3/4, 1)$ let γ have a positive smooth bump of area $1/2$.

Then from the formula above we verify that δ has the following properties:

(a) $\delta = \alpha$ on $[0, 3/4]$. In particular, $\delta = t$ on $[0, 1/4]$ and $\delta(3/4) = 1$.

(b) We have

$$\begin{aligned} \delta(1) &= \int_{3/4}^1 (\beta\dot{\gamma} - \gamma\dot{\beta})\, dt + 1 \\ &= \beta\gamma|_{3/4}^1 - 2\int_{3/4}^1 \gamma\dot{\beta}\, dt + 1 = -2(1/2) + 1 = 0 \end{aligned}$$

and $\delta = 0$ on $[1, \infty)$.

Once we have shown that there is Q satisfying $\ddot{w} = Qw$, we will have produced w with

$$w = \begin{bmatrix} 1 & t-1 \\ 0 & 0 \end{bmatrix}$$

for all $t \geq 1$, which will be the desired example.

Now note that the signs of α, β, γ, δ are such that $\alpha\delta - \beta\gamma > 0$ on $(0, 1)$, so on the open interval Q will exist. Also, since the second derivative of w vanishes on $(0, 1/4]$, we know that Q is identically zero on there, and so Q may be extended smoothly to zero on $(-\infty, 0]$. On $[3/4, 1)$, we have

$$\begin{aligned} Q^{ab} &= \begin{bmatrix} \alpha\ddot{\beta} - \beta\ddot{\alpha} & \ddot{\alpha}\delta - \ddot{\beta}\gamma \\ \alpha\ddot{\delta} - \beta\ddot{\gamma} & \ddot{\gamma}\delta - \gamma\ddot{\delta} \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\gamma}\delta - \gamma\ddot{\delta} \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\gamma}\delta - \gamma\frac{d}{dt}(\beta\dot{\gamma} - \gamma\dot{\beta}) \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\gamma}\delta - \gamma(\beta\ddot{\gamma} - \gamma\ddot{\beta}) \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\gamma}\delta - \gamma\beta\ddot{\gamma} \end{bmatrix} (\alpha\delta - \beta\gamma)^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \ddot{\gamma} \end{bmatrix} \end{aligned}$$

which clearly extends smoothly $[1, \infty)$.

12. Final comments. The theory developed in this paper keeps track of the net number of timelike minus spacelike conjugate points. It is evident from the examples presented in the last section that only this difference is stable, and so any attempt to identify the numbers of timelike, spacelike and null conjugate points separately would require a reckoning sensitive to fine analytic details of $Q^a{}_b$. It is not impossible that such a theory exists. If we take $Q^a{}_b$ to be analytic, then when conjugate points appear to annihilate they really move into the complex. Is it possible to develop an index theorem that, say, relates the number of timelike conjugate points to spectral data and information about the complex conjugate points? If so, what is the difference between taking $Q^a{}_b$ to be C^∞ versus analytic? (Related to this latter, what is the inverse scattering theory for pseudo-self-adjoint operators?)

In the positive-definite Morse index theory, a remarkable corollary of the main theorem is that the number of points conjugate to t_0 on $[t_0, t_1]$ is the same as the number of points conjugate to t_1 on the same interval. (This can be easily pictured by a direct homotopy argument in the Lagrange Grassmannian, too.) In the present situation, however, we know that the conjugate index from t_0 to t_1 is the same as that from t_1 to t_0, but we do not know whether the individual counts of timelike, spacelike and null conjugate points are the same from either endpoint. It seems unlikely, but so far a concrete example has eluded us.

We have not in our treatment needed to derive an explicit differential equation for the curve $L(t)$. This can be done in a straightforward way; the result, if expressed in a canonical chart, is a matrix Riccati equation. As is well-known, scalar Riccati equations have only simple poles as singularities [8], and this remains true for the matrix Riccati equations derived when η is definite. However, when η is indefinite, more exotic singularities can occur, and examples of these are readily derived from the examples in the previous section. In the simplest case the Riccati equation is

$$\dot{L}^a{}_b = -L^a{}_c L^c{}_b + Q^a{}_b$$

where $L^a{}_b = \dot{w}^a_j (w^{-1})^j_b$. The singularities of $L^a{}_b$ are easily analyzed from the explicit formulae for w^a_j.

In the positive-definite case, we can see from the Riccati equation and the fact that all its singularities are isolated that it is possible to recover the function $Q^a{}_b$ from the curve L in Λ. This is no longer true in the indefinite case. In fact, in the example constructed in the last section, we may clearly add to $Q^a{}_b$ any function of the form $f(t)l^a l_b$, where l^a is the vector whose components are

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and f is a smooth function which vanishes for $t \in [0, 1]$. On the other hand, in "generic" situations, the function $Q^a{}_b$ is determined uniquely by the Riccati equation.

References

[1] J. K. Beem and P. E. Ehrlich, *Cut points, conjugate points and Lorentzian comparison theorems*, Math. Proc. Cambridge Philos. Soc., **86** (1979), 365–384.

[2] ——, *A Morse index theorem for null geodesics*, Duke Math. J., **46** (1979), 561–569.

[3] ——, *Global Lorentzian Geometry*, Marcel Dekker, 1981.

[4] R. Bott, *The stable homotopy of the classical groups*, Ann. of Math., **70** (1959), 313–337.

[5] N. Dunford and J. T. Schwartz, *Linear Operators*, Pure and Applied Mathematics, volume VII, Wiley Interscience, 1976-71.

[6] H. Friedrich, *Asymptotic structure of space-time*, in Recent advances in general relativity: essays in honor of E. T. Newman, Birkhäuser, 1992.

[7] I. Gohberg, P. Lancaster and L. Rodman, *Matrices and Indefinite Scalar Products*, Operator Theory: Advances and Applications, vol. 8, Birkhäuser (1983).

[8] E. L. Ince, *Ordinary Differential Equations*, Longmans, Green and Co., 1927.

[9] P. Libermann and C.-M. Marle, *Symplectic Geometry and Analytical Mechanics*, D. Reidel, 1987.

[10] M. Morse, *A generalization of the Sturm separation and comparison theorems in n-space*, Math. Ann., **103** (1930), 52–69.

[11] ——, *The calculus of variations in the large*, Amer. Math. Soc. Colloq. Publ., vol. 18, Amer. Math. Soc., Providence, RI, 1934.

[12] R. Penrose, *Techniques of differential topology in relativity*, Regional Conference Series in Applied Math., vol. 7, SIAM, Philadelphia, PA, 1972.

[13] F. J. Tipler, C. J. S., Clarke and G. F. R. Ellis, *Singularities and horizons—a review article*, in General relativity and gravitation: one hundred years after the birth of Albert Einstein, vol. 2, Plenum Press 1980, pp. 97–206.

[14] F. Treves, *Introduction to Pseudodifferential and Fourier Integral Operators,* vol. 2, Plenum Press, 1982.

Received May 18, 1992 and in revised form August 6, 1992.

University of Missouri
Columbia, MO 65211