TOKYO J. MATH.
VOL. 21, NO. 1, 1998

The Maslov Index: a Functional Analytical Definition and the Spectral Flow Formula

Bernhelm BOOSS-BAVNBEK and Kenro FURUTANI

Roskilde University and Science University of Tokyo

(Communicated by T. Suzuki)

Dedicated to Professor Takeshi Hirai on his 60th birthday

Abstract. We give a functional analytical definition of the Maslov index for continuous curves in the Fredholm-Lagrangian Grassmannian. Our definition does not require assumptions either at the endpoints or at the crossings of the curve with the Maslov cycle. We demonstrate an application of our definition by developing the symplectic geometry of self-adjoint extensions of unbounded symmetric operators. We discuss continuous variations of the form $A_D + C_t$, where A_D is a fixed self-adjoint unbounded Fredholm operator and $\{C_t\}$ a family of bounded self-adjoint operators. We extend the definition of the spectral flow to such families of unbounded operators in a purely functional analytical way. We then prove that the spectral flow is equal to the Maslov index of the corresponding family of abstract Cauchy data spaces.

Introduction.

In this paper we give an elementary, purely functional analytical proof of the *Spectral Flow Formula*. We show that the spectral flow of a continuous curve of self-adjoint (unbounded) Fredholm operators is equal to the Maslov index of the corresponding curve of abstract Cauchy data.

Formulas of this type were first proved by Floer, [14]. He examined a concrete example arising from the symplectic action integral in Hamiltonian mechanics. A general formula was then proved by Yoshida, [33] establishing the equality of the spectral flow and the Maslov index for a family of Dirac operators on a closed 3-dimensional manifold. He applied this identification to the study of 3-dimensional manifold topology. More recently Nicolaescu, [25], [26], [27] gave a generalization of the Spectral Flow Formula for families of Dirac operators on higher dimensional manifolds by assuming invertibility at the endpoints of the path. There are many papers discussing these topics under such assumptions, see e.g. Bunke [9], Cappell, Lee, and Miller, [11], Furutani and Otsuki, [28], [16], Kirk and Klassen, [19], [20], [21], [22].

The purpose of this paper is to provide a purely functional analytical proof of the Spectral Flow Formula that clarifies which of these assumptions are essential. We ignore

Received July 2, 1996
Revised January 9, 1997

all features connected solely with geometric formulations of the problem. To do this, we must clarify the functional analytical meaning of cutting a manifold by a hypersurface. Here this will mean, first, that a self-adjoint operator (typically a Dirac operator) is given over a compact manifold without boundary. A splitting of the manifold is then a process which yields a pair of closed symmetric operators with suitable properties. This is naturally understood as an inverse procedure to the classical von Neumann theory of self-adjoint extensions.

We proceed as follows. In Section 1 we develop the real and complex functional analysis of infinite-dimensional Lagrangians and give a new definition of the Maslov index. Our Maslov index is defined for any continuous path and does not require any deformations of the path, or assumptions at endpoints and crossings. It is applicable to both the finite-dimensional and the infinite-dimensional cases. In the case of cycles, our definition gives the usual Maslov index. For paths, it gives the additivity of the Maslov index under catenation and also its homotopy invariance.

In Section 2 we determine the precise difference between our definition of the Maslov index and that of Robbin and Salamon, who defined the Maslov index for *smooth* curves of Lagrangians which have *regular* crossings with the Maslov cycle (see [31]).

In Section 3 we construct a specific continuous curve of Fredholm pairs of Lagrangians. More precisely, we first define the space β of abstract boundary values of a fixed closed, symmetric operator A in a real Hilbert space to be the quotient space of the maximal and minimal domain of A. Then we equip β with a symplectic structure. Usually, the Cauchy data space for a differential operator over a compact manifold with boundary is defined as the L_2-closure of the space of sufficiently differentiable solutions restricted to the boundary (see [7], Chapter 13 which establishes also the Lagrangian property of the Cauchy data spaces for Dirac operators). We provide a more algebraic argument by exploiting the symplectic structure of β. From this we immediately derive the Lagrangian property of the (abstract) Cauchy data spaces. This property implies that the Cauchy data spaces are closed subspaces of β and vary continuously for a continuous variation $A+C_t$ of A where $\{C_t\}$ is a family of bounded self-adjoint operators. Here 'continuous' refers to the operator norm.

We study operator families which have the following properties. They are continuous families of unbounded closed symmetric operators which all have a self-adjoint extension with compact resolvent. None of the operators has inner solutions. All the operators have the same domain and differ only by a self-adjoint bounded operator.

There are three situations in which such families naturally arise. First, consider a Dirac operator acting on the sections of a fixed Clifford module bundle E over a closed manifold M with fixed Riemannian structure. Varying the connection in E yields a family of Dirac operators. These are closed symmetric operators in $L_2(M; E)$. Their self-adjoint extensions A_t are uniquely determined and have domain equal to the Sobolev space $\mathscr{H}^1(M; E)$. The operators can be naturally seen as perturbations by bundle homomorphisms (i.e. by particularly simple bounded operators) of a densely defined, closed symmetric operator in $L_2(M; E)$. Second, consider a family of Dirac operators on a codimension 0 submanifold M_+ with boundary. The domain $D \subset L_2(M_+; E_{|M_+})$ is specified once we impose a fixed global elliptic boundary condition, but vary the connection. Third, consider a continuous variation of the boundary condition in the Grassmannian of generalized Atiyah-Patodi-Singer projections. If we apply a suitable unitary transformation (as explained by Douglas and Wojciechowski, [13], Formula A10), this case will be reduced to the preceding second case (see also Booss-Bavnbek and Wojciechowski, [8], Section 3).

In Section 4 we give a rigorous definition of the spectral flow for the aforementioned class of continuous families of *unbounded* Fredholm operators. Our goal for the general unbounded case is to obtain a spectral flow which is a homotopy invariant. Recall that the spectral flow is a homotopy invariant for paths of bounded self-adjoint Fredholm operators. Therefore it can be related to certain well-known topological phenomena. It can, for example, be related to the fact that the non-trivial component of the space of *bounded* self-adjoint Fredholm operators is a classifying space for the functor $\mathbf{KR}^{-7}$ in the real case, and for the functor K^{-1} in the complex case (see [5]).

The original definition of the spectral flow was given by Atiyah, Patodi and Singer in [4]. They brought the graph of the spectrum of the family into 'general position' by deformation and counted the number of intersections with $y=0$. But this approach is only meaningful for loops or loop-like curves; for example, operator curves with periodic spectrum or with invertible endpoints. Still this approach is very useful for concrete calculations, if the crossings are smooth and regular.

We proceed in two steps to get a precise definition which is independent of any deformations made. First, we apply Phillips' purely functional analytical definition of the spectral flow for a continuous curve (not necessarily a loop) of *bounded* Fredholm operators (see [30]). Phillips' definition needs no assumptions about the zero eigenvalues and does not require any deformation of the family into 'general position'. More precisely, on each small interval $[t_1, t_2]$ there is a bound $a>0$ such that a does not belong to the spectrum of any A_t and such that only finitely many eigenvalues of A_t belong to the interval $[-a, a]$. We count the number of eigenvalues of A_t which belong to the interval $[0, a]$ for $t=t_2$ and $t=t_1$, and take the difference between these two integers. Then the spectral flow is the sum of these differences for a sufficiently fine partition of the interval $[0, 1]$. In this way, we do not need to count the crossings and avoid any ambiguities. The operators at the endpoints do not have to be invertible and the movement of the eigenvalues do not have to strictly increase or decrease when passing zero.

The second step is to discuss the spectral flow of continuous families of unbounded self-adjoint Fredholm operators of the form $A_t := A_D + C_t$ discussed in Section 3. Clearly, any unbounded self-adjoint Fredholm operator A_t can be transformed into a bounded self-adjoint Fredholm operator by the transformation

$$A_t \mapsto A_t\sqrt{\mathrm{Id}+A_t^2}^{-1}. \tag{0.1}$$

Here $\sqrt{\mathrm{Id}+A_t^2}^{-1}$ denotes the unique positive definite square root of the positive definite operator $(\mathrm{Id}+A_t^2)^{-1}$.

We define the spectral flow of the family $\{A_t\}$ directly on the operator level as the spectral flow of the transformed family. The transformation is, however, *not* continuous on the whole space of unbounded self-adjoint Fredholm operators. B. Fuglede gave a counterexample in [15]. We prove that (0.1) transforms a continuous path $\{A_t = A_D + C_t\}$ of the aforementioned form into a continuous path of bounded self-adjoint Fredholm operators. In other words, we prove the continuity of the *combined* transformation

$$C \mapsto A_D + C \mapsto (A_D + C)(\mathrm{Id} + (A_D + C)^2)^{-1/2}.$$

We shall also assign another continuous curve $\{\tilde{A}_t\}$ of bounded self-adjoint Fredholm operators to the curve $\{A_t\}$. We show that this curve gives the same spectral flow as the curve transformed according to (0.1).

In Section 5 we prove our main result (Theorem 5.1):

THE SPECTRAL FLOW FORMULA. *Let A be a closed symmetric operator in a real Hilbert space H with domain D_m and let $\{C_t\}_{t\in I}$ be a continuous family of bounded self-adjoint operators on H. We assume that*

1. *the operator A has a self-adjoint extension A_D with compact resolvent;*
2. *there exists a positive constant a such that*

$$D_m \cap \ker(A^* + C_t - s) = \{0\}$$

for any s with $|s| < a$ and any $t \in [0, 1]$.

Then we have

$$\mathrm{sf}(\{A_D + C_t\}) = \mu(\{\gamma(\ker(A^* + C_t)\}, \gamma(D)), \tag{0.2}$$

where γ denotes the projection of the domain D_M of A^* onto the symplectic space $\beta = D_M/D_m$ of abstract boundary values.

We construct a two-parameter family of operators and apply the relation formula between the Robbin and Salamon definition and our definition of the Maslov index to analytic segments of these derived families. Then the Spectral Flow Formula follows by a homotopy argument.

ACKNOWLEDGMENTS. We thank R. Nest for helping us to prove Proposition 4.5 and B. Fuglede, G. Grubb, B. Ørsted, N. Otsuki, and K. P. Wojciechowski for stimulating discussions. We also thank the referee for helping us improve the presentation of this paper. Our work was supported by the Danish Science Research Council, grants no. 9501870 and no. 9503564.

1. The Grassmannian of Lagrangian Fredholm pairs.

1.1. Symplectic functional analysis. We fix the following notation.

Let $(\mathscr{H}, \langle\cdot,\cdot\rangle, \omega)$ be a fixed symplectic, separable real Hilbert space and let $J\colon \mathscr{H} \to \mathscr{H}$ denote the corresponding almost complex structure defined by $\omega(x, y) = \langle Jx, y\rangle$ with $J^2 = -\mathrm{Id}$, ${}^tJ = -J$, and $\langle Jx, Jy\rangle = \langle x, y\rangle$. Here ${}^tJ$ denotes the transpose of J with regard to the (real) inner product $\langle\cdot,\cdot\rangle$. Let $\mathscr{L} = \mathscr{L}(\mathscr{H})$ denote the set of all Lagrangian subspaces of $\mathscr{H}$ (i.e. $\Lambda = (J\Lambda)^\perp$). It is naturally identified with the space $\mathscr{C}$ of self-adjoint involutions of $\mathscr{H}$ which anti-commute with J. The correspondence is given by

$$\mathscr{L} \ni \Lambda \mapsto C := 2P_\Lambda - \mathrm{Id} \in \mathscr{C} \quad \text{and} \quad \mathscr{C} \ni C \mapsto \Lambda := \{x \in \mathscr{H} \mid Cx = x\},$$

where P_Λ denotes the orthogonal projection onto Λ. We topologize $\mathscr{L}$ by the topology of $\mathscr{C}$ as a subset in the space $\mathscr{B}(\mathscr{H})$ of bounded operators on $\mathscr{H}$. The space $\mathscr{L}$ is contractible and therefore not topologically interesting. To get something topologically meaningful, we fix a Lagrangian subspace Λ_0.

DEFINITION 1.1. (a) The *Fredholm-Lagrangian Grassmannian* of a real symplectic Hilbert space $\mathscr{H}$ at a fixed Lagrangian subspace Λ_0 is defined as

$$\mathscr{F}\mathscr{L}_{\Lambda_0} := \{\Lambda \in \mathscr{L} \mid (\Lambda, \Lambda_0) \text{ Fredholm pair}\}$$

(i.e. $\dim \Lambda \cap \Lambda_0 < \infty$, $\Lambda + \Lambda_0$ closed and codim $\Lambda + \Lambda_0 < \infty$).

(b) The *Maslov cycle* of Λ_0 in $\mathscr{H}$ is defined as

$$\mathscr{M}_{\Lambda_0} := \mathscr{F}\mathscr{L}_{\Lambda_0} \setminus \mathscr{F}\mathscr{L}^{(0)}_{\Lambda_0},$$

where $\mathscr{F}\mathscr{L}^{(0)}_{\Lambda_0}$ denotes the subset of Lagrangians intersecting Λ_0 transversally, i.e. $\Lambda \cap \Lambda_0 = \{0\}$.

NOTE. The Fredholm-Lagrangian Grassmannian in an infinite-dimensional symplectic Hilbert space was first considered in Swanson, [32].

Using the almost complex structure J, we consider the space $\mathscr{H}$ as a *complex Hilbert space*. We denote it by the same letter $\mathscr{H}$. The complex inner product of $\mathscr{H}$ is given by

$$\langle x, y\rangle_{\mathbf{C}} := \langle x, y\rangle - \sqrt{-1}\,\omega(x, y).$$

Let $\mathscr{U}(\mathscr{H})$ denote the group of all unitary operators on $\mathscr{H}$ and $\mathscr{U}_c(\mathscr{H})$ denote the subgroup of operators of the form $\mathrm{Id} + K$, where K is a compact operator. Then $\mathscr{U}_c(\mathscr{H})$ acts transitively on $\mathscr{F}\mathscr{L}_{\Lambda_0}$ in a natural way. Let

$$\begin{aligned} \rho\colon \mathscr{U}_c(\mathscr{H}) &\longrightarrow \mathscr{F}\mathscr{L}_{\Lambda_0} \\ U &\longmapsto U(\Lambda_0^\perp) \end{aligned}$$

denote the mapping defined by this action. This mapping is the projection of the principal fibre bundle $\mathcal{U}_c(\mathcal{H})$ onto its base space $\mathcal{FL}_{\Lambda_0}$ (see Swanson, [32], Lemma 3).

Since $\mathcal{H}$ can be considered as the complexification $\Lambda_0 \otimes \mathbf{C} \cong \Lambda_0 + J\Lambda_0 = \mathcal{H}$ of Λ_0, we attain a *complex conjugation* by

$$\begin{array}{ccc} z = x \otimes 1 + y \otimes \sqrt{-1} & \longmapsto & x \otimes 1 - y \otimes \sqrt{-1} =: \bar{z} \\ \| & & \| \\ x + Jy & & x - Jy \end{array}$$

where $x, y \in \Lambda_0$. Let $A \in \mathcal{B}(\mathcal{H})$. We let $\bar{A}$ denote the bounded operator on $\mathcal{H}$ given by $\bar{A}(z) := \overline{A(\bar{z})}$ and denote ${}^T A := \bar{A}^*$. Notice that in difference to the real transpose ${}^t A$, the new conjugate ${}^T A$ belongs to the category of complex operators and is defined with respect to the fixed Λ_0.

Let $\Lambda \in \mathcal{FL}_{\Lambda_0}$ be given. We choose an operator $U \in \mathcal{U}_c(\mathcal{H})$ such that $\Lambda = U(\Lambda_0^\perp)$. From the definition of the conjugate ${}^T U$ we obtain the following isomorphy of complex subspaces of $\Lambda_0 \otimes \mathbf{C} \cong \mathcal{H}$:

$$\ker(U\,{}^T U + \mathrm{Id}) \cong \Lambda \cap \Lambda_0 + J(\Lambda \cap \Lambda_0) \cong (\Lambda \cap \Lambda_0) \otimes \mathbf{C}\,. \tag{1.1}$$

Equation (1.1) proves:

Lemma 1.2. *For any $\Lambda \in \mathcal{FL}_{\Lambda_0}$ and any $U \in \mathcal{U}_c(\mathcal{H})$ with $\Lambda = (\Lambda_0^\perp)$, we have*

$$\dim_{\mathbf{R}}(\Lambda \cap \Lambda_0) = \dim_{\mathbf{C}} \ker(U\,{}^T U + \mathrm{Id})\,.$$

Our functional analytical definition of the Maslov index builds upon this property of the operator $U\,{}^T U$.

1.2. The complex Fredholm-Lagrangian Grassmannian and the symmetric generator. Above we considered the complex Hilbert space consisting of the points of $\mathcal{H}$. Now we consider the *complexification* $\mathcal{H} \otimes \mathbf{C}$. This space splits into a direct sum of the two eigenspaces E_+, E_- of $J \otimes \mathrm{Id}$ for the eigenvalues $\pm\sqrt{-1}$. We define the set of *complex Lagrangian subspaces* of $\mathcal{H} \otimes \mathbf{C}$ by

$$\mathcal{L}^{\mathbf{C}} := \{L \subset \mathcal{H} \otimes \mathbf{C} \mid L + (J \otimes \mathrm{Id})(L) = \mathcal{H} \otimes \mathbf{C} \text{ and } \langle L, (J \otimes \mathrm{Id})L \rangle^{\mathbf{C}} = 0\}\,,$$

where $\langle \cdot, \cdot \rangle^{\mathbf{C}}$ denotes the Hermitian inner product in $\mathcal{H} \otimes \mathbf{C}$. We obtain a natural embedding of $\mathcal{L}$ in $\mathcal{L}^{\mathbf{C}}$ given by

$$\mathcal{L} \ni \Lambda \longmapsto \tau(\Lambda) := \Lambda \otimes \mathbf{C} \in \mathcal{L}^{\mathbf{C}}\,.$$

Let $\mathcal{G}$ denote the group of all unitary operators which commute with $J \otimes \mathrm{Id}$, i.e. which keep $\mathsf{E}_\pm$ invariant. Hence $\mathcal{G}$ is isomorphic to $\mathcal{U}(\mathsf{E}_-) \times \mathcal{U}(\mathsf{E}_+)$. Clearly $\mathcal{G}$ acts on $\mathcal{L}^{\mathbf{C}}$. Let $\mathcal{G}_c$ denote the subgroup of operators of the form $\mathrm{Id} + K$, where K is a compact operator on $\mathcal{H} \otimes \mathbf{C}$. Then also this group splits into

$$\mathcal{G}_c \cong \mathcal{U}_c(\mathsf{E}_-) \times \mathcal{U}_c(\mathsf{E}_+)\,.$$

We define the *complex Fredholm-Lagrangian Grassmannian* by

$$\mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}} := \{L \in \mathscr{L}^{\mathbf{C}} \mid (L, \Lambda_0 \otimes \mathbf{C}) \text{ Fredholm pair}\}\ .$$

The group $\mathscr{G}_c$ acts transitively on $\mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$. Let

$$\begin{aligned} \rho^{\mathbf{C}}: \mathscr{G}_c &\longrightarrow \mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}} \\ g &\longmapsto g(\Lambda_0^{\perp} \otimes \mathbf{C}) \end{aligned}$$

denote the map defined by this action. We obtain a commutative diagram

$$\begin{array}{ccc} \mathscr{U}_c(\mathscr{H}) & \xrightarrow{\tilde{\tau}} & \mathscr{G}_c \\ \downarrow \rho & & \downarrow \rho^{\mathbf{C}} \\ \mathscr{F}\mathscr{L}_{\Lambda_0} & \xrightarrow{\tau} & \mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}\ , \end{array}$$

where $\tilde{\tau}$ denotes the complexification $U \mapsto U \otimes \mathrm{Id} = \begin{pmatrix} \bar{U} & 0 \\ 0 & U \end{pmatrix}$. Here we identify $\mathscr{U}_c(\mathscr{H})$ with $\mathscr{U}_c(\mathsf{E}_{\pm})$ by the (complex and anti-linear) isomorphisms

$$\begin{aligned} \mathscr{H} &\xrightarrow{\cong} \mathsf{E}_{\pm} \\ z &\longmapsto \frac{z \otimes 1 \mp J(z) \otimes \sqrt{-1}}{\sqrt{2}}\ , \end{aligned}$$

so that $\bar{U}$ operates on E_- and U on E_+ in the preceding matrix.

Now we construct an isomorphism between $\mathscr{U}_c(\mathscr{H})$ and $\mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$. Let Φ be the mapping

$$\begin{aligned} \Phi: \mathscr{U}_c(\mathscr{H}) &\longrightarrow \mathscr{G}_c \\ U &\longmapsto \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & U \end{pmatrix}, \end{aligned}$$

where Id operates on E_- and U is considered as an operator $\mathsf{E}_+ \to \mathsf{E}_+$. We split

$$\mathscr{G}_c = \left\{\begin{pmatrix} \mathrm{Id} & 0 \\ 0 & U \end{pmatrix}\right\} \times \left\{\begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix}\right\} = \left\{\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}\right\},$$

and see that the range of Φ and the second factor of $\mathscr{G}_c$ intersect only at the identity. We also see that the right inverse of Φ is given by

$$\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \longmapsto VU^{-1}\ ,$$

where U^{-1} is considered as an operator from E_+ to E_+ by successively identifying $\mathsf{E}_- \cong \mathscr{H}_{\mathbf{C}} \cong \mathsf{E}_+$.

The preceding facts prove

PROPOSITION 1.3. *For any real symplectic, separable Hilbert space $\mathscr{H}$ with fixed Lagrangian Λ_0, we have a homeomorphism*

$$\rho^{\mathbf{C}} \circ \Phi : \mathscr{U}_c(\mathscr{H}) \xrightarrow{\sim} \mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0 \otimes \mathbf{C}} .$$

Note that the inverse mapping

$$\Psi := (\rho^{\mathbf{C}} \circ \Phi)^{-1} : \mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0 \otimes \mathbf{C}} \to \mathscr{U}_c(\mathscr{H})$$

is explicitly given on the range of τ by

$$\Lambda \otimes \mathbf{C} \mapsto U\bar{U}^{-1} = U^T U \qquad \text{where} \quad U \in \mathscr{U}_c(\mathscr{H}) \quad \text{with } \Lambda = U(\Lambda_0^{\perp}) .$$

Hence we can introduce our key operator for the direct functional analytical definition of the Maslov index:

DEFINITION 1.4. For any $\Lambda \in \mathscr{F}\mathscr{L}_{\Lambda_0}$, we define the *complex symmetric generator* of Λ (with regard to Λ_0) by

$$W_\Lambda := \Psi(\Lambda \otimes \mathbf{C}) \overset{\text{Prop. 1.3}}{=\!=\!=} U^T U \in \mathscr{U}_c(\mathscr{H}) .$$

NOTE. The operator $W_\Lambda = U^T U$ is invariantly defined by Λ. This was already found by Leray, [23], Lemma 2.1, by direct calculation in the *real* Grassmannian. In our context, the invariance of W_Λ is just a geometric property of the *complex* Grassmannian. Consider the principal fibre bundle given by the action of the group $\mathscr{G}_c$ on $\mathscr{F}\mathscr{L}^{\mathbf{C}}_{\Lambda_0 \otimes \mathbf{C}}$. The bundle has a global section provided by Ψ. That is the reason for the invariance of W_Λ. Moreover, we can describe the set of all operators which arise as complex generators W_Λ of (real) Fredholm-Lagrangians Λ. It is exactly the subset of all $W \in \mathscr{U}_c(\mathscr{H})$ with $W = {}^T W$.

1.3. A proper functional analytical definition of the Maslov index for arbitrary paths. Let $\{\Lambda(t)\}_{t \in I}$ with $I = [0, 1]$ be a continuous path in $\mathscr{F}\mathscr{L}_{\Lambda_0}$. Then the family $\{W_{\Lambda(t)}\}_{t \in I}$ of unitary operators on $\mathscr{H}$ is also a continuous family in the operator norm. To define the Maslov index we proceed in a similar way as Phillips did when he gave a direct definition of the spectral flow of a continuous path of self-adjoint, bounded Fredholm operators (for details see [30] or below Section 4).

To define the spectral flow one deals with an operator family with spectral oscillations on the real line around zero. To define the Maslov index we define an operator family with spectral oscillations on the unit circle around $e^{i\pi}$. In both cases we want to count the net number of eigenvalues, counted with multiplicities, which pass through a fixed gauge in the positive direction.

In general, it is not possible to lock the oscillations of the eigenvalues into an interval $[-a, a]$ (or into an arc between $e^{i(\pi - a)}$ and $e^{i(\pi + a)}$) so that no eigenvalues can leak through the boundary $\pm a$ when the parameter runs from 0 to 1. But Phillips observed

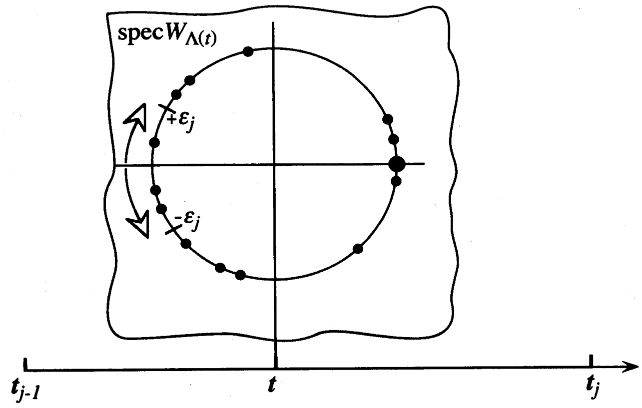


FIGURE 1. Horizontal and vertical spacing of the spectrum of the complex symmetric generator $W_{\Lambda(t)}$

that the strategy works locally and can be patched together. We shall use a similar approach to define the Maslov index.

We choose a partition $\{0 = t_0 < t_1 < \cdots < t_N = 1\}$ of the interval and positive numbers $0 < \varepsilon_j < \pi$, $j = 1, \cdots, N$, such that

$$\ker(W_{\Lambda(t)} - e^{i(\pi \pm \varepsilon_j)}) = \{0\} \tag{1.2}$$

for $t_{j-1} \leq t \leq t_j$ (see Figure 1). Here we use the fact that $W_{\Lambda(t)} - e^{i\pi}$ is a Fredholm operator. This is clear since $W_{\Lambda(t)}$ is unitary with eigenvalues on the unit circle and since $W_{\Lambda(t)} - \mathrm{Id}$ is compact with discrete eigenvalues and with 0 as the only accumulation point.

If the equation (1.2) is satisfied at a point t for an ε_j, it will be satisfied also in near neighbouring points with the same bound ε_j. In such a way we can construct a finite number of intervals $[t_{j-1}, t_j]$ which are a partition of the interval $[0, 1]$ and which satisfy equation (1.2) for suitable bounds ε_j.

We consider a continuous curve in the Fredholm-Lagrangian and admit that the curve intersects with the Maslov cycle $\mathcal{M}_{\Lambda_0}$ for an infinite number of parameters. Hence, there is in general not a concrete intersection number to define the Maslov index. We now count a number which characterizes the oscillation and which coincides with the intersection number of the curve with the Maslov cycle, if the curve is in a 'general position'.

DEFINITION 1.5. Let $\Lambda = \{\Lambda(t)\}_{t \in I}$ be a continuous path in the Fredholm-Lagrangian Grassmannian $\mathcal{FL}_{\Lambda_0}$ of a real symplectic Hilbert space $\mathcal{H}$ at a fixed Lagrangian Λ_0. We define the *Maslov index* by

$$\mu(\Lambda) = \mu(\Lambda; \Lambda_0) := \sum_{j=1}^{N} k(t_j, \varepsilon_j) - k(t_{j-1}, \varepsilon_j)$$

with

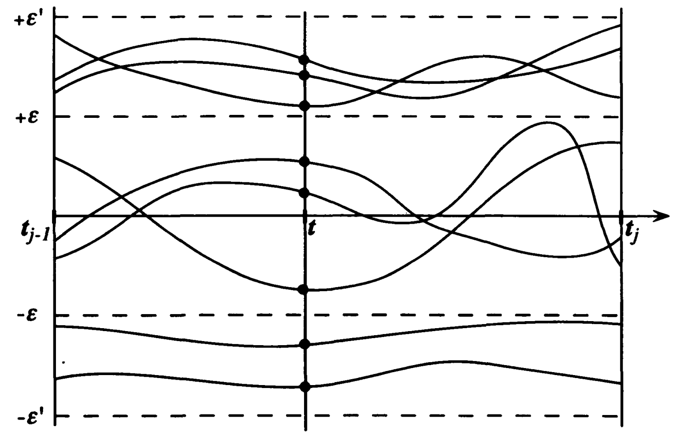


FIGURE 2. The locking of the eigenvalues between ε and ε'

$$k(t, \varepsilon_j) := \sum_{0 \leq \theta < \varepsilon_j} \dim \ker(W_{\Lambda(t)} - e^{i(\pi+\theta)}) \qquad \text{for} \quad t_{j-1} \leq t \leq t_j ,$$

where the horizontal and vertical spacing $(t_0, \cdots, t_N)$, $(\varepsilon_1, \cdots, \varepsilon_N)$ is chosen as in (1.2).

Notice that our definition of the Maslov index does not depend on the choice of the horizontal or vertical spacing. This is not only because of the continuity of the eigenvalues but also because the role of ε_j, chosen as in (1.2), is to lock the eigenvalues for $t \in [t_{j-1}, t_j]$ in the interval $[-\varepsilon, \varepsilon]$. Choosing any other such ε' will also lock the eigenvalues between ε and ε', see Figure 2. This proves the claimed independence of our definition of the Maslov index of the choice of the horizontal or vertical spacing.

All Robbin and Salamon's 'axioms' for the Maslov index (see [31], Theorem 2.3, and, in similar form, Cappell, Lee, and Miller, [10]) follows at once from our construction of the Maslov index. We emphasize the following properties.

THEOREM 1.6. (I) *The Maslov index is well defined for* homotopy *classes of paths with fixed endpoints. In particular, the Maslov index is invariant under re-parametrization of paths.*

(II) *The Maslov index is additive under* catenation, *i.e.*

$$\mu(\Lambda_1 * \Lambda_2) = \mu(\Lambda_1) + \mu(\Lambda_2) ,$$

where $\{\Lambda_1(t)\}$, $\{\Lambda_2(t)\}$ *are continuous paths with* $\Lambda_1(1) = \Lambda_2(0)$ *and*

$$(\Lambda_1 * \Lambda_2)(t) := \begin{cases} \Lambda_1(2t) & 0 \leq t \leq 1/2 \\ \Lambda_2(2t-1) & 1/2 < t \leq 1 . \end{cases}$$

(III) *The Maslov index is* natural *under the action of the group* $\mathrm{Sp}(\mathscr{H})$ *of symplectic automorphisms of* $\mathscr{H}$.

(IV) *The Maslov index vanishes for paths which stay in one stratum* $\mathscr{FL}^{(k)}_{\Lambda_0}$ *of the stratified space* $\mathscr{FL}_{\Lambda_0} = \bigcup_{k=0}^{\infty} \mathscr{FL}^{(k)}_{\Lambda_0}$, *i.e. if* $\dim \Lambda(t) \cap \Lambda_0 = k$ *for one* $k \geq 0$ *and all* $t \in I$.

To define the Maslov index, we embedded the real Fredholm Grassmannian $\mathcal{FL}_{\Lambda_0}$ in the complex Grassmannian $\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$. The inclusion was given by complexification. Clearly, there are many more Lagrangian curves in $\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$ than those coming from $\mathcal{FL}_{\Lambda_0}$. Also the *complex Maslov cycle* defined by

$$\mathcal{M}_{\Lambda_0\otimes\mathbf{C}}:=\{L\in\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}} \mid L\cap(\Lambda_0\otimes\mathbf{C})\neq\{0\}\}$$

is substantially larger than the real Maslov cycle as defined before. We can generalize Lemma 1.2 and characterize the complex Maslov cycle by the property that for all its L, the operator $\Psi(L)$ has eigenvalue -1. This leads to a genuinely *complex* version of Definition 1.5:

DEFINITION 1.7. For any continuous family $\{L(t)\}\in\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$, we define the *complex Maslov index* by

$$\mu^{\mathbf{C}}(\{L(t)\}):=\sum_{j=1}^{N} k(t_j,\varepsilon_j)-k(t_{j-1},\varepsilon_j)\,,$$

where we replace the operator $W_{\Lambda(t)}$ by the operator $\Psi(L(t))$ in the definition of the multiplicities $k(t.,\varepsilon.)$.

Notice that Theorem 1.6 remains valid in the complex case. Furthermore, we see at once that $\mu^{\mathbf{C}}$ is the intersection number of the family $L(t)$ with the complex Maslov cycle, if $\{L(t)\}$ is in a 'general position'. It is not difficult to derive the following formula:

PROPOSITON 1.8. *Let $\{\Lambda(t)\}\in\mathcal{FL}_{\Lambda_0}$ and $\{L(t)\}\in\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$ be two continuous families which have the same endpoints. They are homotopic in $\mathcal{FL}^{\mathbf{C}}_{\Lambda_0\otimes\mathbf{C}}$, if and only if*

$$\mu(\{\Lambda(t)\})=\mu^{\mathbf{C}}(\{L(t)\})\,.$$

REMARK 1.9. We notice that $\mu(\{\Lambda(t)\})$ is the winding number of the closed curve $\{\det W_{\Lambda(t)}\}_{t\in S^1}$ for loops, i.e. for $\Lambda(0)=\Lambda(1)$, and for finite-dimensional $\mathcal{H}$. This is the original definition of the Maslov index as explained in Arnold, [3]. Arnold's definition can be transferred to infinite-dimensional inductive limits. It was generalized by Swanson, [32] to cycles in the Fredholm-Lagrangian Grassmannian. Similarly, we get for arbitrary complex Lagrangian loops that $\mu^{\mathbf{C}}(\{L(t)\})$ is the winding number of $\{\det\Psi(L(t))\}_{t\in S^1}$.

2. The relation between the differential and functional analytical definition of the Maslov index.

2.1. Review of the differential definition. Here we assume that the symplectic vector space $\mathcal{H}$ is finite-dimensional. Notice, though, that the definition of the Maslov index given by Robbin and Salamon, [31], can be extended immediately to the infinite-dimensional case following Swanson [32], Theorem 1.2, where the differentiable structure for $\mathcal{FL}_{\Lambda_0}$ was defined (see also Nicolaescu [26]). In our proof of the Spectral

Flow Formula (in Section 5) we shall apply the Robbin and Salamon definition in the infinite-dimensional form to a simple analytic family. Restricting ourselves to the finite-dimensional case makes the presentation more easy, since we can then identify $\mathcal{FL}_{\Lambda_0}$ with $\mathcal{L}$ and $\mathcal{U}_c(\mathcal{H})$ with $\mathcal{U}(\mathcal{H})$. However, we still fix one Lagrangian subspace Λ_0 to specify the Maslov cycle.

Roughly speaking, Robbin and Salamon's *differential* approach for defining the Maslov index is based on three observations:

(i) Tangent vectors $(\Lambda, \dot{\Lambda}) \in T_\Lambda(\mathcal{L})$ can be regarded as symmetric bilinear forms $Q_{(\Lambda,\dot{\Lambda})}$ on Λ in a natural way.

(ii) For a C^1 (or smooth) curve $\{\Lambda(t)\}_{t\in I}$ of Lagrangian subspaces there is a natural way to distinguish the special case when the curve has only regular crossings with the Maslov cycle $\mathcal{M}_{\Lambda_0}$. Here *regular crossing* at $t\in I$ means that the symmetric bilinear form

$$Q_{(\Lambda(t),\dot{\Lambda}(t))\,|\,\Lambda(t)\cap\Lambda_0}: (\Lambda(t)\cap\Lambda_0)\times(\Lambda(t)\cap\Lambda_0)\to\mathbf{R}$$

is non-singular.

(iii) Since regular crossings are isolated, one can define a number by adding the signatures and possible corrections at the ends of the path. The number is the Robbin-Salamon (differential) Maslov index. Clearly, it will remain unchanged by any further deformation within the class of smooth curves with only regular crossings with the Maslov cycle. (It seems that there is no clear argument in the literature how one can establish that the number will neither change under deformations passing through the class of continuous curves with only regular crossings. See also Remark 2.2 below.)

To explain observation (i), we set for a $(\Lambda, \dot{\Lambda}) \in T_\Lambda(\mathcal{L})$ and $x, y \in \Lambda$:

$$Q_{(\Lambda,\dot{\Lambda})}(x,y) := \frac{d}{ds}\,\omega(x, B_s y)_{|s=0}\,, \tag{2.1}$$

where the family $\{B_s\colon \Lambda\to J(\Lambda)\}_{|s|\ll 1}$ of linear maps is chosen in such a way that its graph $\Lambda(s) := \{x + B_s x \mid x\in\Lambda\}$ becomes a C^1-curve through Λ at $s=0$ with $(d/ds)\Lambda(s)_{|s=0} = \dot{\Lambda}$.

We assume that $\{\Lambda_t\}_{t\in I}$ is a C^1-curve with only regular crossings with the Maslov cycle $\mathcal{M}_{\Lambda_0}$. From [31] we recall Robbin and Salamon's definition of the (differential) Maslov index by

$$\begin{aligned}\mu^{RS}(\Lambda;\Lambda_0) := {}& \tfrac{1}{2}\,\mathrm{sign}\, Q_{(\Lambda(0),\dot{\Lambda}(0))\,|\,\Lambda(0)\cap\Lambda_0} \\ &+ \sum_{0<t<1} \mathrm{sign}\, Q_{(\Lambda(t),\dot{\Lambda}(t))\,|\,\Lambda(t)\cap\Lambda_0} + \tfrac{1}{2}\,\mathrm{sign}\, Q_{(\Lambda(1),\dot{\Lambda}(1))\,|\,\Lambda(1)\cap\Lambda_0}\,.\end{aligned} \tag{2.2}$$

2.2. The relation between the differential and functional analytical definition for smooth paths. Let us assume that the curve $\{\Lambda(t)\}_{t\in I}$ is of C^2-class and has only regular crossings with the Maslov cycle $\mathcal{M}_{\Lambda_0}$. We show that our functional analytical definition of the Maslov index coincides with Robbin and Salamon's differential definition except for the corrections at the endpoints of the path. These corrections sometimes cause the differential Maslov index to become a half-integer, while our definition always provides an integer (as wanted for **Z**-valued homotopy invariants). More precisely we have:

THEOREM 2.1. *Under the preceding conditions of C^2-differentiability and regular crossings we have*

$$\mu=\mu^{RS}-\frac{k_{t=0}}{2}+\frac{k_{t=1}}{2}, \tag{2.3}$$

where k_t denotes the crossing dimension $\dim\Lambda(t)\cap\Lambda_0$.

REMARK 2.2. Formula (2.3) also establishes the homotopy invariance of Robbin and Salamon's definition by avoiding the delicate argument of deformations within the class of curves with only regular crossings.

PROOF. First we show that our and Robbin and Salamon's definitions coincide in suitable small intervals. We relate the eigenvalues of our symmetric unitary generator $W_{\Lambda(t)}$ on both sides of $e^{i\pi}$ with the positive and negative eigenvalues of the quadratic form $Q_{(\Lambda(t),\dot\Lambda(t))\,|\,\Lambda(t)\cap\Lambda_0}$. Later we shall add over a partition of the unit interval I into small intervals and compare the eigenvalues at the endpoints of I.

Step 1: We consider a small neighbourhood of a point t_0, $0<t_0<1$, where $\Lambda(t_0)\cap\Lambda_0\neq\{0\}$. Let $U_t\in\mathcal{U}(\mathcal{H})$ be a curve of unitary transformations such that $U_t(J\Lambda_0)=\Lambda(t)$ for $|t-t_0|\ll 1$. If we write $U_t=X_t+\sqrt{-1}\,Y_t$, we can express the quadratic form on the variable space $\Lambda(t)$ of (2.1) as a quadratic form on the fixed space Λ_0 by substituting $x=U_{t_0}Ju$ and $y=U_{t_0}Jv$ with $u, v\in\Lambda_0$. As observed already by Robbin and Salamon, the coordinate change yields

$$Q_{(\Lambda(t_0),\dot\Lambda(t_0))}(U_{t_0}Ju,\ U_{t_0}Jv)=\langle\dot Y_{t_0}(u),\ X_{t_0}(v)\rangle-\langle\dot X_{t_0}(u),\ Y_{t_0}(v)\rangle\,. \tag{2.4}$$

Step 2: Note that $\Psi(\Lambda(t)\otimes\mathbf{C})=U_t{}^TU_t=:W_t$ is our $W_{\Lambda(t)}$ of Definition 1.4. Writing $U_t=U_{t_0}e^{iA_t}$ and $W_t=W_{t_0}e^{iS_t}$ with self-adjoint A_t and S_t and $S_{t_0}=0$ yields

$$Q_{(\Lambda(t_0),\dot\Lambda(t_0))}(-Y_{t_0}u+JX_{t_0}u,\ -Y_{t_0}v+JX_{t_0}v)=\langle\dot a_{t_0}(u),v\rangle\,, \tag{2.5}$$

where $A_t=a_t+ib_t$. Also, we have the unitary equivalence

$${}^TU_{t_0}\dot S_{t_0}=2\dot a_{t_0}{}^TU_{t_0}\,. \tag{2.6}$$

Equations (2.5) and (2.6) imply that $\operatorname{sign}Q_{(\Lambda(t_0),\dot\Lambda(t_0))\,|\,\Lambda(t_0)\cap\Lambda_0}$ coincides with the signatures of $\dot a_{t_0}$ and $\dot S_{t_0}$ on corresponding subspaces of Λ_0. Notice that the imaginary part ib_{t_0} disappears in the signature formula for very good reasons, namely because the choice of U_t is non-unique.

Step 3: Now we relate the signature of the quadratic form at t_0 to the curve of eigenvalues of W_t for $|t-t_0|\ll 1$. We assume that

$$\dim_{\mathbf{R}}(\Lambda(t_0)\cap\Lambda_0)=\dim\ker(W_{t_0}-e^{i\pi})=k>0\,.$$

Now we lock the eigenvalues of W_t at $t=t_0$ ('vertically') by $\varepsilon>0$ such that

$$\ker(W_{t_0}-e^{i(\pi+\theta)})=\{0\}\qquad\text{for}\quad 0<|\theta|\leq\varepsilon$$

and ('horizontally') by $\delta>0$ such that $\ker(W_t-e^{i(\pi\pm\varepsilon)})$ remains equal to $\{0\}$. Hence

$$\sum_{|\theta|\leq\varepsilon}\dim\ker(W_t-e^{i(\pi+\theta)})=k$$

for $|t-t_0|<\delta$.

Let

$$0<\lambda_1\leq\cdots\leq\lambda_p\quad\text{and}\quad 0>\mu_1\geq\cdots\geq\mu_q$$

denote the eigenvalues of $\dot{S}_{t_0|\Lambda(t_0)\cap\Lambda_0}$. We have no vanishing eigenvalues since the crossing is assumed to be regular, hence $p+q=k$. Assume that $\Lambda(t)$ is of C^2-class. The transformation W_t will have eigenvalues $\{\lambda_l(t)\}$ and $\{\mu_j(t)\}$ for t sufficiently close to t_0, say in the interval $[t_0-\delta,\,t_0+\delta]$. These eigenvalues bifurcate from -1 at t_0 in the following form:

$$\begin{aligned}\lambda_l(t)&=e^{i(\pi+\lambda_l t+O(t^2))}\,,\qquad l=1,\cdots,p\,,\\ \mu_j(t)&=e^{i(\pi+\mu_j t+O(t^2))}\,,\qquad j=1,\cdots,q\,.\end{aligned}$$

It follows that the point t_0, where $\dim_{\mathbf{R}}(\Lambda(t_0)\cap\Lambda_0)>0$, is isolated and that

$$\sum_{0\leq\theta\leq\varepsilon}\dim\ker(W_t-e^{i(\pi+\theta)})=p\quad\text{and}\quad\sum_{-\varepsilon\leq\theta<0}\dim\ker(W_t-e^{i(\pi+\theta)})=q$$

for $t_0<t\leq t_0+\delta$ (and vice versa for $t_0-\delta\leq t<t_0$). Hence

$$\begin{aligned}\mu(\{\Lambda(t)\}_{t_0-\delta\leq t\leq t_0+\delta};\,\Lambda_0)&=k(t_0+\delta,\,\varepsilon)-k(t_0-\delta,\,\varepsilon)=p-q\\ &=\operatorname{sign}Q_{(\Lambda(t_0),\dot{\Lambda}(t_0))|\Lambda(t_0)\cap\Lambda_0}=\mu^{RS}(\{\Lambda(t)\}_{t_0-\delta\leq t\leq t_0+\delta};\,\Lambda_0)\,.\end{aligned}\tag{2.7}$$

Step 4: We still have to compare the counting at the endpoints if the crossings are not transversal. At the left endpoint $t_0=0$ we have $k(0+\delta,\varepsilon)=p$ and $k(0,\varepsilon)=\dim\Lambda(0)\cap\Lambda_0=k=p+q$; hence our definition of the Maslov index contributes with $p-(p+q)=-q$, while Robbin and Salamon's definition contributes with $(p-q)/2=k/2-q$. Similarly, at the right endpoint $t_0=1$, we get $k(1,\varepsilon')-k(1-\delta',\,\varepsilon')=k'-q'=p'$, while Robbin and Salamon get $(p'-q')/2=p'-k'/2$. That explains the error terms in the formula (2.3). We obtain the full proof of our theorem by the additivity under catenation of paths. □

3. An example: the Fredholm Lagrangian of abstract Cauchy data spaces.

3.1. The symplectic space of abstract boundary values. We assume that H is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$; that D_m is a dense subsapce of H; and that A is a closed symmetric operator in H defined on D_m. Let D_M denote the domain of the adjoint operator A^* of A. We have

$$D_M \supset D_m \quad \text{and} \quad A^*{}_{|D_m} = A\,,$$

i.e. A^* is a (closed) extension of A. Let $D_M^{\mathscr{G}}$ and $D_m^{\mathscr{G}}$ denote the corresponding Hilbert spaces, equipped with the inner product coming from the graph norm

$$\langle x, y\rangle_{\mathscr{G}} := \langle x, y\rangle + \langle A^*x, A^*y\rangle \qquad \text{for} \quad x, y \in D_M\,.$$

Then $D_m^{\mathscr{G}}$ is a closed subspace of $D_M^{\mathscr{G}}$. We denote by $\|x\|_{\mathscr{G}} := \sqrt{\langle x, x\rangle_{\mathscr{G}}} = \sqrt{\langle x, x\rangle + \langle A^*x, A^*x\rangle}$. Set $\beta := D_M^{\mathscr{G}}/D_m^{\mathscr{G}}$ with the canonical projection

$$\begin{aligned} \gamma : D_M^{\mathscr{G}} &\longrightarrow \quad \beta \\ x &\longmapsto [x] = x + D_m \end{aligned}$$

and with the quotient norm

$$\|\gamma(x)\|_\beta := \inf_{a \in D_m} \|x + a\|_{\mathscr{G}} \qquad \text{for} \quad \gamma(x) \in \beta\,.$$

We call β the space of *abstract boundary values* and γ the *abstract trace map*. We have a short exact sequence of Hilbert spaces

$$0 \longrightarrow D_m^{\mathscr{G}} \hookrightarrow D_M^{\mathscr{G}} \xrightarrow{\ \gamma\ } \beta \longrightarrow 0 \tag{3.1}$$

which splits with a right inverse j of γ. Then $\tilde{\beta} := j(\beta)$ is a closed subspace of $D_M^{\mathscr{G}}$, characterized by

$$\tilde{\beta} \cong \beta\,, \quad D_M^{\mathscr{G}} = D_m^{\mathscr{G}} \oplus \tilde{\beta} \quad \text{and} \quad D_m^{\mathscr{G}} \perp_{\mathscr{G}} \tilde{\beta}\,. \tag{3.2}$$

More precisely, we have:

LEMMA 3.1. (a) *The space β of abstract boundary values can be represented in $D_M^{\mathscr{G}}$ as the -1-eigenspace of the square $(A^*)^2$ of the (real) symmetric operator A^*:*

$$\tilde{\beta} = \{ y \in D_M \mid A^*y \in D_M \text{ and } (A^*)^2 y = -y \}\,. \tag{3.3}$$

(b) *Let $D \subset D_M^{\mathscr{G}}$ be a subspace with $D_m \subset D$. Then D is closed in $D_M^{\mathscr{G}}$, if and only if $\gamma(D)$ is closed in β.*

PROOF. (a) Let $b \in \beta$, say $b = \gamma(z)$ with $z \in D_M^{\mathscr{G}}$. We split $z = x + y$, where $x \in D_m^{\mathscr{G}}$ and $y \in D_m^{\mathscr{G}\perp}$. Then $y \perp_{\mathscr{G}} D_m^{\mathscr{G}}$ implies that $\langle x, y\rangle + \langle A^*x, A^*y\rangle = 0$. By definition $\langle A^*x, A^*y\rangle = \langle x, A^*A^*y\rangle$. Hence $y \perp_{\mathscr{G}} D_m^{\mathscr{G}}$, if and only if $\langle x, A^*A^*y\rangle = \langle x, -y\rangle$ for all $x \in D_m^{\mathscr{G}}$, i.e. $(A^*)^2 y = -y$.

(b) If D is closed in $D_M^{\mathscr{G}}$, then the factor space $D/D_m^{\mathscr{G}} = \gamma(D)$ is a complete space, hence it is closed in β. □

We introduce a symplectic structure on β by setting

$$\omega([x],[y]) := \langle A^*x, y\rangle - \langle x, A^*y\rangle \quad \text{for} \quad [x],[y]\in\beta. \tag{3.4}$$

PROPOSITION 3.2. *The form ω is a well-defined skew-symmetric bilinear form on $\beta\times\beta$ with the following properties*:

(i) *ω is bounded*;

(ii) *ω is non-degenerate.*

PROOF. (i) is proved as follows, using the elementary algebraic inequality $\sqrt{Bc}+\sqrt{bC}\le\sqrt{b+B}\sqrt{c+C}$ for non-negative reals.

$$\begin{aligned}|\omega([x],[y])| &\le |\langle A^*x, y\rangle| + |\langle x, A^*y\rangle| \le \|A^*x\|\,\|y\| + \|x\|\,\|A^*y\| \\ &\le \sqrt{\|x\|^2+\|A^*x\|^2}\sqrt{\|y\|^2+\|A^*y\|^2} = \|x\|_{\mathscr{G}}\|y\|_{\mathscr{G}},\end{aligned} \tag{3.5}$$

where $x, y\in D_M$ $(=D_M^{\mathscr{G}})$. Hence $|\omega([x],[y])|\le\|[x]\|_\beta\|[y]\|_\beta$.

To prove (ii) we lift ω to the representation $\tilde\beta$ of β in $D_M^{\mathscr{G}}$. Let $\tilde\omega$ denote the form $\tilde\omega(x,y) := \langle A^*x, y\rangle - \langle x, A^*y\rangle$ restricted to $\tilde\beta$, i.e.

$$\tilde\omega(j([x]), j([y])) = \omega([x],[y]) \quad \text{for all} \quad [x],[y]\in\beta.$$

Notice that

$$A^*(\tilde\beta)\subset\tilde\beta \quad\text{and}\quad (A^*)^2 = -\mathrm{Id} \quad\text{on } \tilde\beta, \quad \text{and} \tag{3.6}$$

$$\tilde\omega(x,y) = \langle A^*x, y\rangle_{\mathscr{G}}. \tag{3.7}$$

From (3.6) and (3.7) we see that the mapping

$$\begin{aligned}\tau:\ \beta &\longrightarrow \beta^* \\ [x] &\longmapsto \tau_{[x]}([y]) := \omega([x],[y])\end{aligned}$$

is an isomorphism of the Hilbert space β onto its dual β^*; hence (ii) is proved. □

We can characterize various types of extensions of the fixed symmetric, closed operator A by the corresponding properties of the domains and by the abstract boundary values (see also [1], [2]).

LEMMA 3.3. *Let D be a subspace of D_M, which contains D_m. Then the extension $A_D := A^*|_D$*

(a) *is closed (as an operator in H), if and only if $\gamma(D)$ is closed (in β)*;

(b) *the extension is self-adjoint, if and only if $\gamma(D)$ is a Lagrangian subspace of β*; *and*

(c) *it has compact resolvent, if and only if the inclusion $D^{\mathscr{G}} \hookrightarrow H$ is compact, where $D^{\mathscr{G}}$ denotes the domain D equipped with the graph norm.*

PROOF. (a) is just a reformulation of Lemma 3.1b; (b) and (c) are immediate from the definition. □

3.2. Lagrangian property of Cauchy data spaces. As first suggested by Bojarski,

[6], the concept of Cauchy data spaces is fundamental to any systematic study of splitting formulas for spectral invariants. This motivates the following definition in our abstract setting:

DEFINITION 3.4. Let

$$\mathsf{S} := \ker A^*$$

denote the solution space of A^*. The space S is closed in the graph norm in $D_M^{\mathscr{G}}$ and also in H. We call $\gamma(\mathsf{S})$ the *Cauchy data space* of A^*.

All the arguments in this section will assume the following:

ASSUMPTION 1. There exists a self-adjoint Fredholm extension

$$A_D := A^*{}_{|D}$$

defined on a domain D with $D_m \subset D \subset D_M$. In particular, $\gamma(D)$ is a Lagrangian subspace in β.

Assuming the existence of a Fredholm extension in our abstract setting corresponds to the ellipticity condition in the concrete setting. We shall exploit the following list of Fredholm properties:

⟨1⟩ By definition, $\ker A_D = D \cap \mathsf{S}$ is finite-dimensional.

⟨2⟩ We have a short exact sequence

$$0 \longrightarrow D_m \cap \mathsf{S} \hookrightarrow D \cap \mathsf{S} \xrightarrow{\gamma_{|D \cap \mathsf{S}}} \gamma(D \cap \mathsf{S}) \longrightarrow 0\,,$$

which yields $D \cap \mathsf{S} \cong D_m \cap \mathsf{S} \oplus \gamma(D \cap \mathsf{S})$.

⟨3⟩ Clearly $\gamma(D \cap \mathsf{S}) \subset \gamma(D) \cap \gamma(\mathsf{S})$; in fact the spaces are equal, since $\gamma(x) = \gamma(s)$ for $x \in D$ and $s \in \mathsf{S}$ implies $x - s \in D_m$. Hence $s \in D$.

⟨4⟩ range $A_D = A^*(D)$ is closed in H and $\dim H/\text{range} A_D < +\infty$, so $A^*(D_M)$ is also closed in H.

⟨5⟩ $\ker A = D_m \cap \mathsf{S} = A^*(D_M)^\perp$, with the orthogonal complement taken in H.

⟨6⟩ $\ker A_D = (\text{range } A_D)^\perp$, with the orthogonal complement taken in H.

Assumption 1 leads to the following proposition which is the main result of this section.

PROPOSITION 3.5. *The Cauchy data space $\gamma(\mathsf{S})$ is a closed, Lagrangian subspace of β and belongs to the Fredholm-Lagrangian Grassmannian $\mathscr{FL}_{\Lambda_0}$ at $\Lambda_0 := \gamma(D)$.*

It is an astonishing aspect of symplectic functional analysis that the proof of the preceding proposition can be kept completely elementary due to the following geometric comparison lemma.

LEMMA 3.6. *Let (V, ω) be a real symplectic Hilbert space with a fixed Lagrangian subspace Λ_0 and an isotropic subspace Λ. Then Λ is Lagrangian, if*

$$\Lambda_0 \cap \bar{\Lambda} = \{0\} \quad \textit{and} \quad \Lambda_0 + \Lambda = V .$$

PROOF. Let $x \in \bar{\Lambda}$, say $x = x_0 + x_1$ with $x_0 \in \Lambda_0$ and $x_1 \in \Lambda \subset \bar{\Lambda}$. Hence $x_0 = x - x_1 \in \Lambda_0 \cap \bar{\Lambda}$, which must vanish, so $x = x_1 \in \Lambda$. This proves $\bar{\Lambda} = \Lambda$.

To prove the Lagrangian property of Λ, we take $x \in \Lambda^\circ$ and write it as $x = x_0 + x_1$ as before. Since $\Lambda \subset \Lambda^\circ$, we get

$$x_0 = x - x_1 \in \Lambda_0 \cap \Lambda^\circ = \Lambda_0^\circ \cap \Lambda^\circ = (\Lambda_0 + \Lambda)^\circ = V^\circ = \{0\} ,$$

hence $x = x_1 \in \Lambda$. □

PROOF OF PROPOSITION 3.5. *Step* 1: Let $[x], [y] \in \gamma(\mathrm{S})$. Then

$$\omega([x], [y]) = \langle A^* x, y \rangle - \langle x, A^* y \rangle = 0 ,$$

hence $\gamma(\mathrm{S})$ isotropic.

Step 2: Now we consider the sequence of continuous mappings

$$D_M^{\mathscr{G}} \xrightarrow{A^*} \text{range } A^* \xrightarrow{\pi} \text{range } A^* / \text{range } A_D .$$

Hence

$$D + \mathrm{S} = \{ x \in D_M^{\mathscr{G}} \mid A^* x \in \text{range } A_D \} = \ker \pi \circ A^*$$

must be closed, and we have a Hilbert space isomorphism

$$D_M^{\mathscr{G}} / (D + \mathrm{S}) \overset{\rightarrow}{\cong} \text{range } A^* / \text{range } A_D . \tag{3.8}$$

Moreover, $\gamma(D + \mathrm{S})$ is closed in β by Lemma 3.1b and coincides with $\gamma(D) + \gamma(\mathrm{S})$. From the closedness of $\gamma(D) + \gamma(\mathrm{S})$, we get

$$(\gamma(D) + \gamma(\mathrm{S}))^\circ = (\gamma(D) + \overline{\gamma(\mathrm{S})})^\circ = \gamma(D)^\circ \cap \overline{\gamma(\mathrm{S})}^\circ . \tag{3.9}$$

Step 3: Since $\gamma(\mathrm{S})$ is isotropic, also $\overline{\gamma(\mathrm{S})}$ is isotropic. Recall that $\gamma(D)$ is Lagrangian. This yields

$$\gamma(D)^\circ \cap \overline{\gamma(\mathrm{S})}^\circ \supset \gamma(D) \cap \overline{\gamma(\mathrm{S})} \supset \gamma(D) \cap \gamma(\mathrm{S}) ,$$

hence, with (3.9)

$$(\gamma(D) + \gamma(\mathrm{S}))^\circ \supset \gamma(D) \cap \gamma(\mathrm{S}) . \tag{3.10}$$

Step 4: Now we exploit the Fredholm properties and get

$$D_m \cap \mathrm{S} \oplus \gamma(D) \cap \gamma(\mathrm{S}) \overset{\langle 2 \rangle, \langle 3 \rangle}{\cong} D \cap \mathrm{S} \overset{\langle 1 \rangle}{=} \ker A_D \tag{3.11}$$

$$\overset{\langle 6 \rangle}{=} (\text{range } A_D)^\perp = H / \text{range } A_D \cong H / \text{range } A^* \oplus \text{range } A^* / \text{range } A_D .$$

Since $D_m \cap \mathrm{S} \overset{\langle 5\rangle}{\cong} H/\operatorname{range} A^*$, this yields

$$(3.12)\qquad \gamma(D)\cap\gamma(\mathrm{S})\cong \operatorname{range} A^*/\operatorname{range} A_D \overset{(3.8)}{\cong} D_M/(D+\mathrm{S})$$
$$\cong \beta/\gamma(D+\mathrm{S})=\beta/(\gamma(D)+\gamma(\mathrm{S}))\,.$$

Moreover, for any closed subspace L in β, we have

$$\beta/L^\circ \cong \beta^*/\tau_\omega(L)\cong \beta^*/\tau_E(L)\cong \beta/L^\perp\,,$$

where the isomorphisms $\tau_E, \tau_\omega : \beta\to\beta^*$ are given by $\tau_E([x])[y] := \langle [x], [y]\rangle_\beta$ and $\tau_\omega([x])[y] := \omega([x], [y])$. Hence

$$(3.13)\qquad \dim \beta/\gamma(D)+\gamma(\mathrm{S}) = \dim(\gamma(D)+\gamma(\mathrm{S}))^\circ\,.$$

Combined with formulas (3.10) and (3.12) this yields

$$(3.14)\qquad (\gamma(D)+\gamma(\mathrm{S}))^\circ = \gamma(D)\cap\overline{\gamma(\mathrm{S})} = \gamma(D)\cap\gamma(\mathrm{S})\,.$$

Step 5: Set $\mu := \Lambda\cap\gamma(\mathrm{S})$. Since μ is finite-dimensional, it is also closed. (3.14) yields $\mu\subset\mu^\circ=\gamma(D)+\gamma(\mathrm{S})$, i.e. μ is isotropic. Hence, in the reduced symplectic vector space μ°/μ, we have

$$(3.15)\qquad \gamma(D)/\mu\cap\overline{\gamma(\mathrm{S})/\mu}=\{0\} \quad\text{and}\quad \gamma(D)/\mu+\gamma(\mathrm{S})/\mu=\mu^\circ/\mu\,.$$

Clearly, $\gamma(D)/\mu$ is Lagrangian in the factor space; hence we can apply Lemma 3.6 and get that $\gamma(\mathrm{S})/\mu$ is Lagrangian in μ°/μ, hence $\gamma(\mathrm{S})$ Lagrangian in μ° and in β.

From Formula (3.13) we see that $\gamma(\mathrm{S})$ and $\gamma(D)$ form a Fredholm pair. □

COROLLARY 3.7. *Let Λ be a Lagrangian subspace in β. Then $(\Lambda, \gamma(\mathrm{S}))$ is a Fredholm pair, if and only if $A_{\gamma^{-1}(\Lambda)} := A^*|_{\gamma^{-1}(\Lambda)}$ is a (self-adjoint) Fredholm operator. We then have*

$$\operatorname{index} A_{\gamma^{-1}(\Lambda)} = \mathbf{i}(\Lambda, \gamma(\mathrm{S})) = \dim \Lambda\cap\gamma(\mathrm{S}) - \operatorname{codim}(\Lambda+\gamma(\mathrm{S})) = 0\,.$$

3.3. The continuity of the Cauchy data spaces. We shall investigate the Cauchy data spaces of operator families of the form $\{A^*+C_t\}_{t\in I}$, where A is a closed symmetric, densely defined operator in a Hilbert space H which satisfies suitable additional assumptions. We assume that $\{C_t\}_{t\in I}$ is a continuous family (with respect to the operator norm) of bounded self-adjoint operators. Here the parameter t runs in the interval $I=[0, 1]$.

We define the space of abstract boundary values and the abstract trace map $\gamma : D_M \to D_M^{\mathscr{G}}/D_m^{\mathscr{G}} = \beta$ as before. Notice that the vector spaces β and the mapping γ are fixed even in the family situation; but, given by the graph of A^*+C_t, the inner product $\langle\cdot,\cdot\rangle_t^{\mathscr{G}}$ for $D_M^{\mathscr{G}}$ and β varies with varying parameter t. Hence the splitting $j_t : \beta\to D_M^{\mathscr{G}}$ varies and so does the representation of β as subspace $j_t(\beta)=\tilde\beta_t$ in D_M; yet all norms are uniformly equivalent with respect to $t\in[0, 1]$, and equivalent to the norm defined just by A^*. Hence, in the following we fix the inner product defined by A^*. Moreover, the symplectic structure of the space β, as defined in Formula (3.4), does not depend on the parameter t.

We sharpen our previous Assumption 1 by demanding the existence of a domain D with $D_m \subset D \subset D_M$, such that $A_D := A^*{}_{|D}$ has compact resolvent. Hence, the operators $A_D + C_t$ are Fredholm operators for that fixed D and all $t \in I$.

ASSUMPTION 2. We shall assume the non-existence of inner solutions for all operators $A^* + C_t$, i.e.

$$D_m \cap \mathsf{S}_t = \{0\} \qquad \text{for all} \quad t \in [0, 1] .$$

NOTE. The non-existence of inner solutions ('unique continuation property') is not generally valid for elliptic differential operators, but it is established for Dirac operators (see e.g. [7], Chapter 8).

THEOREM 3.8. *Under the preceding assumptions (existence of a self-adjoint extension A_D with compact resolvent and non-existence of inner solutions), the spaces $\gamma(\mathsf{S}_t)$ of Cauchy data of a continuous family $\{A^* + C_t\}_{t \in I}$ vary continuously.*

NOTE. As usual, we define the continuous dependence of a family of subspaces of a Hilbert space on a parameter by the continuity of the corresponding orthogonal projections.

PROOF. To prove the continuity, we need only to consider the local situation at $t=0$. First we show that $\{\mathsf{S}_t\}_{t \in I}$ is a continuous family of subspaces of $D_M^{\mathscr{G}}$; then we show that $\gamma(\mathsf{S}_t)$ is a continuous family in β.

We consider the bounded operator

$$\begin{array}{rcl} F_t : D_M^{\mathscr{G}} & \longrightarrow & H \oplus \mathsf{S}_0 \\ x & \longmapsto & ((A^* + C_t)(x), P_0 x) \end{array}$$

where $P_0 : H \to \mathsf{S}_0$ denotes the orthogonal projection of H onto the subspace S_0, which is closed in $D_M^{\mathscr{G}}$ and in H.

Clearly, F_0 is injective: $F_0(x) = 0$ implies $x \in \mathsf{S}_0$ and $x = P_0 x = 0$. The operator F_0 is also surjective: Since $A^* + C_0$ has no inner solutions, we have $\ker A + C_0 = D_m \cap \mathsf{S}_0 \cong \operatorname{coker} A^* + C_0$ which shows that the operator $A^* + C_0$ is surjective. Let $y \in H$ and $x \in \mathsf{S}_0$ and choose z with $(A^* + C_0)z = y$. Let $w := P_0(z) - x \in \mathsf{S}_0$. Then $F_0(z - w) = ((A^* + C_0)(z - w) = y, P_0(z - w) = x)$. This proves that F_0 is an isomorphism.

Then all operators F_t are isomorphisms for small $t \geq 0$, since F_t is a continuous family of operators. We define

$$\varphi_t := F_t^{-1} \circ F_0 : D_M^{\mathscr{G}} \cong D_M^{\mathscr{G}} \qquad \text{for} \quad t \text{ small} .$$

We see that

$$\varphi_t(\mathsf{S}_0) = \mathsf{S}_t , \tag{3.16}$$

since each $z \in \varphi_t(\mathrm{S}_0)$ implies $F_t(z) = 0 + P_0(z)$. Hence $(A^* + C_t)z = 0$; vice versa, each $z \in \mathrm{S}_t$ can be written in the form $F_t^{-1} F_0(y)$ with $y := P_0(z)$.

From (3.16) we get that

$$\{P_t := \varphi_t P_0 \varphi_t^{-1} : D_M^{\mathscr{G}} \to S_t\}$$

is a continuous family of projections onto the solution spaces S_t. The projections are not necessarily orthogonal, but can be orthogonalized and remain continuous in t like in [7], Lemma 12.8.

Now we must show that $\{\gamma(\mathrm{S}_t)\}$ is a continuous family in β. This is not proved by the formula $\gamma(\mathrm{S}_t) = \gamma(\varphi_t(\mathrm{S}_0))$ alone. We must modify the endomorphism φ_t of $D_M^{\mathscr{G}}$ in such a way that it keeps the subspace D_m invariant. To do that we notice that $D_m + \mathrm{S}_0$ is closed in $D_M^{\mathscr{G}}$. We define a continuous family of mappings by

$$\begin{array}{rcccccl} \psi_t : D_M^{\mathscr{G}} = & D_m + \mathrm{S}_0 & + & (D_m + \mathrm{S}_0)^{\perp} & \longrightarrow & D_M^{\mathscr{G}} \\ & x + s & + & y & \longmapsto & x + \varphi_t(s) + y \end{array}$$

with $\psi_0 = \mathrm{Id}$. Hence ψ_t isomorphism for $t \ll 1$, and $\psi_t(D_m) = D_m$ for such small t. Hence we obtain a continuous family of mappings $\{\tilde{\psi}_t : \beta \to \beta\}$ with $\tilde{\psi}_t(\gamma(\mathrm{S}_0)) = \gamma(\mathrm{S}_t)$. From that we obtain a continuous family of projections as above. □

Remark 3.9. From the preceding arguments it also follows that the Cauchy data spaces form a differentiable family, if $\{C_t\}$ is a differentiable family.

4. The spectral flow for families of self-adjoint (unbounded) Fredholm operators.

4.1. Phillips' definition for continuous bounded families. Let H be a real separable Hilbert space and let $\hat{\mathscr{F}}$ denote the space of bounded self-adjoint Fredholm operators from H to H. It is well known that $\hat{\mathscr{F}}$ consists of three connected components (in the operator norm)

$$\hat{\mathscr{F}} = \hat{\mathscr{F}}_- \cup \hat{\mathscr{F}}_+ \cup \hat{\mathscr{F}}_*,$$

namely the contractible spaces of essentially negative and essentially positive operators and the topologically non-trivial component of operators with essential spectrum on both sides of the real line.

Definition 4.1 [J. Phillips, 1995]. For any arbitrary continuous path $A : [0, 1] \ni t \mapsto A_t \in \hat{\mathscr{F}}$, we define the *spectral flow* by

$$\mathrm{sf}(A) := \sum_{j=1}^{N} k(t_j, \varepsilon_j) - k(t_{j-1}, \varepsilon_j)$$

with

$$k(t, \varepsilon_j) := \sum_{0 \le \theta < \varepsilon_j} \dim \ker(A_t - \theta) \qquad \text{for} \quad t_{j-1} \le t \le t_j,$$

where the horizontal and vertical spacings $(t_0, \cdots, t_N)$, $(\varepsilon_1, \cdots, \varepsilon_N)$ are chosen so that

$$\ker(A_t - \varepsilon_j) = \{0\} \quad \text{and} \quad \sum_{|\theta| < \varepsilon_j} \dim\ker(A_t - \theta) < \infty \tag{4.1}$$

for $t_{j-1} \leq t \leq t_j$ and $0 \leq |\theta| < \varepsilon_j$.

It is possible to choose a vertical and horizontal spacing which satisfies (4.1), since the spectrum of a self-adjoint bounded Fredholm operator changes continuously with the operator and the zero eigenvalue is discrete and of finite multiplicity. After Definition 1.5 we already mentioned Phillips' argument why the definition does not depend on the choice of the horizontal and vertical spacing.

We list the following properties of the spectral flow to emphasize formal similarities with the Maslov index (see Theorem 1.6):

THEOREM 4.2. (I′) *The spectral flow is well defined for homotopy classes of paths with fixed endpoints and it distinguishes the homotopy classes. In particular, it is invariant under re-parametrization of paths.*

(II′) *The spectral flow is additive under catenation, i.e.*

$$\mathrm{sf}(A * B) = \mathrm{sf}(A) + \mathrm{sf}(B)\,,$$

where $\{A_t\}$, $\{B_t\}$ *are continuous paths with* $A_1 = B_0$ *and*

$$(A * B)_t := \begin{cases} A_{2t} & 0 \leq t \leq 1/2 \\ B_{2t-1} & 1/2 < t \leq 1\,. \end{cases}$$

(III′) *The spectral flow is invariant under the adjoint action of the full orthogonal group* $\mathcal{O}(H)$ *of* H.

(IV′) *The spectral flow vanishes for paths which stay in one (connected) stratum*

$$\hat{\mathcal{F}}_{\#}^{(k)} := \{F \in \hat{\mathcal{F}}_{\#} \mid \dim\ker F = k\}\,, \qquad \# \in \{-, +, *\}$$

of the stratified space $\hat{\mathcal{F}}_{\#} = \bigcup_{k=0}^{\infty} \hat{\mathcal{F}}_{\#}^{(k)}$, *i.e. if* $\dim\ker A_t = k$ *for all* $t \in I$ *and one* $k \geq 0$.

We can discuss the relations with the spectral flow $\mathrm{sf}^{\mathbf{C}}$ of the complex case in exactly the same way as we did in Section 1 for the Maslov index: First we embed the space $\mathcal{F} = \mathcal{F}(H)$ of Fredholm operators, defined on the real Hilbert space H, in the complex Fredholm operator space $\mathcal{F}(H \otimes \mathbf{C})$ with the inclusion given by complexification. Clearly, there are many more paths in $\hat{\mathcal{F}}(H \otimes \mathbf{C})$ than those coming from $\hat{\mathcal{F}}(H)$.

Theorem 4.2 remains valid in the complex case. It is not difficult to derive the following formula:

PROPOSITION 4.3. *Let* $\{A_t\} \in \hat{\mathcal{F}}(H)$ *and* $\{B_t\} \in \hat{\mathcal{F}}(H \otimes \mathbf{C})$ *be two continuous paths which have the same endpoints. They are homotopic in* $\hat{\mathcal{F}}(H \otimes \mathbf{C})$, *if and only if*

$$\mathrm{sf}(\{A_t\}) = \mathrm{sf}^{\mathbf{C}}(\{B_t\})\,.$$

Remark 4.4. In spite of formal similarities between the definition of the spectral flow and of the Maslov index, it must be noted that the spectral flow of a path in finite dimension depends only on the eigenvalues at the endpoints. Consequently, it vanishes for loops. This remains true for paths and loops in the components $\hat{\mathscr{F}}_{\pm}$. The Maslov index, however, depends on the path and not only on the endpoints, even in finite dimension. Hence the spectral flow (counting passages through 0 on the real line) is topologically only interesting when we have an infinite number of eigenvalues (or essential spectrum) on both sides of the real line. The Maslov index (counting passages through $e^{i\pi}$ on the circle) is always topologically interesting. Our understanding of the spectral flow as a quantum type invariant is nourished also by the observation that the spectral flow is defined directly by operators and their eigenvalues and not by 'classical' quantities. Moreover, it demands genuinely infinite-dimensional function spaces. In spite of its coincidence with the Maslov index (see below Section 5), the spectral flow reflects the finer distinction between the components $\hat{\mathscr{F}}_{-}$, $\hat{\mathscr{F}}_{+}$, and $\hat{\mathscr{F}}_{*}$.

4.2. The construction of a continuous curve of bounded operators by the transformation $A \mapsto A(\mathrm{Id}+A^2)^{-1/2}$. Now we consider a path $\{A_D+C_t\}_{t\in I}$ of (unbounded) self-adjoint Fredholm operators in H, where A_D is a fixed (unbounded) self-adjoint operator with compact resolvent, and $\{C_t\}_{t\in I}$ is a continuous path of bounded self-adjoint operators on H.

To define the spectral flow of the family $\{A_D+C_t\}_{t\in I}$, we apply the transformation

$$\begin{array}{rccl} \mathscr{R}\colon & C\hat{\mathscr{F}} & \longrightarrow & \hat{\mathscr{F}} \\ & A & \longmapsto & \mathscr{R}(A):=A\sqrt{\mathrm{Id}+A^2}^{-1}, \end{array} \tag{4.2}$$

where $C\hat{\mathscr{F}}$ denotes the space of (not necessarily bounded) self-adjoint Fredholm operators. We define the convergence in $C\hat{\mathscr{F}}$ by the *gap* metric, i.e. the convergence of the orthogonal projection operators onto the graphs of the Fredholm operators. Cordes and Labrousse showed ([12], Addendum, Theorem 1), that on the subset of all bounded operators, the topology, induced by the gap metric for closed operators, is equivalent to that given by the operator norm.

The transformation $\mathscr{R}$ maps the connected component of $C\hat{\mathscr{F}}$ which contains $\hat{\mathscr{F}}_{*}$ into $\hat{\mathscr{F}}_{*}$. The same holds for $\hat{\mathscr{F}}_{\pm}$. From the Spectral Decomposition Theorem and the Weierstass Approximation Theorem it follows that the mapping $\mathscr{R}$, restricted to $\hat{\mathscr{F}}_{*}$ (or $\hat{\mathscr{F}}_{\pm}$), is continuous and homotopic to the identity map of $\hat{\mathscr{F}}_{*}$ (or $\hat{\mathscr{F}}_{\pm}$), see Atiyah and Singer [5]. However, a counterexample by Fuglede [15] shows that the mapping $\mathscr{R}$ is *not* continuous on the whole space $C\hat{\mathscr{F}}$, nor on the subspace $\mathscr{C}$ of self-adjoint operators with compact resolvent.

We shall show the continuity of the composed map

$$C \mapsto A_D + C \mapsto \mathscr{R}(A_D + C)$$

from $\hat{\mathscr{B}}$ to $\hat{\mathscr{F}}$, where $\hat{\mathscr{B}}$ denotes the space of bounded self-adjoint operators on H.

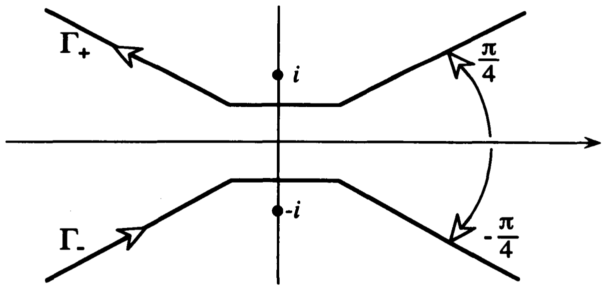


FIGURE 3. The integration path $\Gamma = \Gamma_+ \cup \Gamma_-$

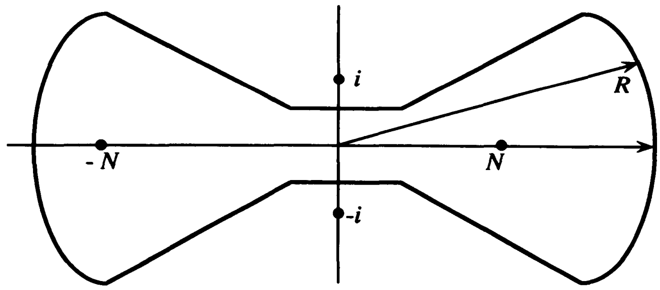


FIGURE 4. The integration path Γ_R for $R \gg N$

PROPOSITION 4.5. *Let S be a self-adjoint operator with spectral decomposition $S = \int_{-\infty}^{\infty} \lambda dE_\lambda$ and set $f_\varepsilon(z) := ze^{-(1/2+\varepsilon)\log(1+z^2)}$ for $\varepsilon > 0$.*[1] *Let $\Gamma = \Gamma_- \cup \Gamma_+$ denote the double cone around the x-axis with opening $(-\pi/4, \pi/4)$ turned off zero by passing inside $\pm i$ (see Figure 3). Then the 'Cauchy integral' converges and defines a bounded operator with*

$$\frac{1}{2\pi i}\int_\Gamma f_\varepsilon(\lambda)(\lambda - S)^{-1} d\lambda = \int_{-\infty}^{\infty} f_\varepsilon(\theta) dE_\theta =: f_\varepsilon(S)\,. \tag{4.3}$$

PROOF (Communicated by R. Nest; see also [28] and [16]). The integral on the right side of (4.3) exists and defines a bounded operator, since the function f_ε is bounded. It follows from the estimate

$$|f_\varepsilon(z)| \sim 1/|z|^{2\varepsilon} \quad \text{as} \quad |z| \to \infty \tag{4.4}$$

that also the integral on the left side of (4.3) is well defined and defines a bounded operator. Therefore, to prove (4.3) it is enough to prove the equality of the two operators on the dense subspace $\bigcup_{N>0} P_N(H)$, where $P_N := \int_{-N}^{N} dE_\theta$.

Let $x \in H$ and $N > 0$. For $R \gg N$, we replace the infinite integration path Γ by the finite closed contour Γ_R as indicated in Figure 4. Then the operator $f_\varepsilon(S)$ takes the form

[1] We fix the branch of $\log(1+z^2)$ for which $-\pi < \arg\log(1+z^2) < \pi$.

$$(4.5)\qquad \int_{-N}^{N} f_\varepsilon(\theta)dE_\theta = \int_{-N}^{N} \frac{1}{2\pi i}\int_{\Gamma_R} \frac{f_\varepsilon(\lambda)}{\lambda-\theta}\,d\lambda dE_\theta$$

$$= \int_{-N}^{N} \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\Gamma_R} \frac{f_\varepsilon(\lambda)}{\lambda-\theta}\,d\lambda dE_\theta = \int_{-N}^{N} \frac{1}{2\pi i}\int_{\Gamma} \frac{f_\varepsilon(\lambda)}{\lambda-\theta}\,d\lambda dE_\theta$$

on the 'compact element' $P_N(x)$. The last equality is proved by applying the estimation (4.4). The last operator applied to $P_N(x)$ yields

$$(4.6)\qquad \frac{1}{2\pi i}\int_\Gamma f_\varepsilon(\lambda)\left(\int_{-N}^{N} \frac{1}{\lambda-\theta}\,dE_\theta\right)(P_N(x))d\lambda$$

$$= \frac{1}{2\pi i}\int_\Gamma f_\varepsilon(\lambda)\left(\int_{-\infty}^{\infty} \frac{1}{\lambda-\theta}\,dE_\theta\right)(P_N(x))d\lambda$$

$$= \frac{1}{2\pi i}\left(\int_\Gamma f_\varepsilon(\lambda)(\lambda-S)^{-1}d\lambda\right)(P_N(x)),$$

which proves equation (4.3). □

REMARK 4.6. The preceding proof must be carried out *after* complexifying H, if the Hilbert space H is real. Anyway, the resulting operator $f_\varepsilon(S)$ remains real so that $f_\varepsilon(S)$ can be considered an operator on H.

The following lemma is a consequence of the preceding proposition.

LEMMA 4.7. *Let S and C be self-adjoint operators with S unbounded and C bounded. Then we have in the operator norm*

$$\|f_\varepsilon(S+C)-f_\varepsilon(S)\| \le c\|C\|,$$

where the constant c does not depend on S, C, or $\varepsilon>0$.

PROOF. By Proposition 4.5 we have

$$(4.7)\qquad f_\varepsilon(S+C)-f_\varepsilon(S) = \frac{1}{2\pi i}\int_\Gamma f_\varepsilon(\lambda)((\lambda-(S+C))^{-1}-(\lambda-S)^{-1})d\lambda$$

$$= \frac{-1}{2\pi i}\int_\Gamma f_\varepsilon(\lambda)(\lambda-(S+C))^{-1}\circ C\circ(\lambda-S)^{-1}d\lambda.$$

Therefore

$$\|f_\varepsilon(S+C)-f_\varepsilon(S)\| \le \frac{1}{2\pi}\int_\Gamma |f_\varepsilon(\lambda)|\frac{1}{|\Im(\lambda)|^2}\|C\|d\lambda = c\|C\|.$$ □

Now we can conclude the main result of this subsection from Proposition 4.5 and Lemma 4.7. We exploit that the limit $\lim_{\varepsilon\to 0} f_\varepsilon(S)$ exists in the strong sense for any self-adjoint operator S with compact resolvent and equals our transformed operator $\mathscr{R}(S) := S(\mathrm{Id}+S^2)^{-1/2}$ by Lebesgue's Convergence Theorem and the Resonance Theorem.

THEOREM 4.8. *Let S be a self-adjoint operator with compact resolvent in a real separable Hilbert space H and let C be a bounded self-adjoint operator. Then the sum $S+C$ also has compact resolvent and is a closed Fredholm operator. We have*

$$\|\mathscr{R}(S+C)-\mathscr{R}(S)\| \leq c\|C\| ,$$

where the constant c does not depend on S or on C.

We shall apply the preceding theorem in the following form:

COROLLARY 4.9. *Curves of self-adjoint (unbounded) Fredholm operators in a separable real Hilbert space of the form $\{A_D+C_t\}_{t\in I}$ are mapped into continuous curves in $\hat{\mathscr{F}}$ by the transformation $\mathscr{R}$ when A_D has compact resolvent and $\{C_t\}_{t\in I}$ is a continuous curve of bounded operators.*

REMARK 4.10. Let

$$\begin{array}{rccc} \mathscr{T}_{A_D}: & \hat{\mathscr{B}} & \longrightarrow & C\hat{\mathscr{F}} \\ & C & \longmapsto & A_D+C \end{array}$$

denote the translation by A_D, mapping bounded self-adjoint operators on H into closed self-adjoint Fredholm operators in H. On $C\hat{\mathscr{F}}$, the gap topology is defined by the metric

$$g(A_1, A_2) := \sqrt{\|R_{A_1}-R_{A_2}\|^2+\|A_1R_{A_1}-A_2R_{A_2}\|^2} ,$$

where $R_A := (\mathrm{Id}+A^2)^{-1}$ (see Cordes and Labrousse, [12] and also Kato, [17]). In Theorem 4.8 we proved that the composition $\mathscr{R}\circ\mathscr{T}_{A_D}$ is continuous.

$$\begin{array}{ccc} \hat{\mathscr{B}} & \xrightarrow{\ \mathscr{T}_{A_D}\ } & C\hat{\mathscr{F}} \\ & {\scriptstyle \mathscr{R}\circ\mathscr{T}_{A_D}} \searrow & \downarrow {\scriptstyle \mathscr{R}} \\ & & \hat{\mathscr{F}} \end{array}$$

Further, we can prove that the translation operator $\mathscr{T}_{A_D}$ is a continuous operator from $\hat{\mathscr{B}}$ onto the subspace $\hat{\mathscr{B}}+A_D \subset C\hat{\mathscr{F}}$. The proof can be carried out along the same lines as the proof of Theorem 4.8 taking the functions $(1+\lambda^2)^{-1}$ and $\lambda(1+\lambda^2)^{-1}$ instead of $f_\varepsilon(\lambda)$. Thus it needs not take a limit $\varepsilon \to 0$.

We define:

DEFINITION 4.11. Let $\{A_t\}_{t\in I}$ be a continuous curve of (unbounded) self-adjoint Fredholm operators of the form $A_t = A_D+C_t$ with the preceding assumptions. Then the *spectral flow* of the continuous, unbounded curve $\{A_t\}_{t\in I}$ is defined by the spectral flow of the continuous, bounded curve $\{\mathscr{R}(A_t)\}_{t\in I}$ in $\hat{\mathscr{F}}$.

REMARK 4.12. The properties listed in Theorem 4.2 for the spectral flow of families of bounded operators remain valid for our class of unbounded operators by the construction of the curve $\{\mathscr{R}(A_t)\}_{t\in I}$.

4.3. An alternative construction of a continuous curve of bounded operators. Phillips' idea was to define the spectral flow of a continuous curve of bounded self-adjoint Fredholm operators by locking curves of eigenvalues piecewise between constant bounds. Now we want to construct a continuous curve $\{\tilde{A}_t\}$ of bounded self-adjoint Fredholm operators little by little from small intervals of the spectrum of our curve $\{A_t = A_D + C_t\}_{t\in I}$ and from the related eigenspaces. We lock the branching of the zero eigen-values between constant bounds. We proceed in a similar way as in the proof of Theorem 2.1 when we determined the relations between the functional analytical and the differential definitions of the Maslov index.

PROPOSITION 4.13. *Let $\{A_t = A_D + C_t\}_{t\in I}$ be a continuous family of self-adjoint Fredholm operators. Then there exist a partition $0 = t_0 < \cdots < t_N = 1$ of the interval and continuous curves $\{A_t^{(j)}\}$ in $\hat{\mathscr{F}}_*$ on each small interval $[t_j, t_{j+1}]$ such that*

1. $\mathrm{spec}_{ess}(A_t^{(j)}) = \{1, -1\}$ *and*

$$\mathrm{spec}(A_t^{(j)}) = \{\mathrm{spec}(A_D + C_t) \cap (-a_j, a_j)\} \cup \{1, -1\}$$

for suitable positive reals $a_1, \cdots, a_N$; and

2. $\ker A_t^{(j)} = \ker(A_D + C_t)$ *for $t_j \le t \le t_{j+1}$.*

PROOF. *Step* 1: To construct the jump curve, we first consider the family $\{A_s\}$ in a neighbourhood $t - \delta(t) \le s \le t + \delta(t)$ of a point $t \in I$, where $\ker A_t = \{0\}$. Then none of the A_s has any eigenvalue in a small vertical interval. Hence there is no contribution to the spectral flow and we can define $A_s^{(t)} := T$ for s in the interval $[t - \delta(t), t + \delta(t)]$. Here $T: H \to H$ denotes an isomorphism which is equal to Id on an infinite-dimensional subspace $L_{(t)}$ and equal to $-\mathrm{Id}$ on the orthogonal, also infinite-dimensional subspace $L'_{(t)}$, with the polarization $(L_{(t)}, L'_{(t)})$ chosen arbitrarily.

Step 2: We consider the family $\{A_s\}$ close to a point t, where we have $\dim \ker A_t > 0$. Let λ_1 denote the smallest positive eigenvalue and μ_1 the largest negative one. We choose a positive real number $a(t) < 1$ with $a(t) < \lambda_1$ and $a(t) < |\mu_1|$ and a $\delta(t) > 0$, such that $a(t), -a(t) \notin \mathrm{spec} A_s$ for $s \in [t - \delta(t), t + \delta(t)]$. For s in this interval, we define

$$P_s^{(t)} := \frac{1}{2\pi i} \int_{|\lambda| = a(t)} (A_s - \lambda)^{-1} d\lambda .$$

It follows that $\mathrm{rank}\, P_s^{(t)} = \dim \ker A_t$. Then the operator $A_s P_s^{(t)}: H \to H$ is bounded because $P_s^{(t)}$ is of finite rank and A_s keeps range $P_s^{(t)}$ invariant for $t - \delta(t) \le s \le t + \delta(t)$.

Step 3: Now we can choose points $\tilde{t}_0 = 0 < \tilde{t}_1 < \cdots < \tilde{t}_N = 1$ in such a way that $\tilde{t}_{j+1} - \delta(\tilde{t}_{j+1}) < \tilde{t}_j + \delta(\tilde{t}_j)$, and then choose points $t_j \in (\tilde{t}_j - \delta(\tilde{t}_j), \tilde{t}_{j-1} + \delta(\tilde{t}_{j-1}))$ with $0 = t_0 = \tilde{t}_0 < \tilde{t}_1 < t_1 < \tilde{t}_2 < \cdots < \tilde{t}_N < t_N = 1$. We set $a_j := a(\tilde{t}_j)$ for $j = 0, \cdots, N$.

Next we choose polarizations of the infinite-dimension

$$(\operatorname{range} P_{t_j}^{(\tilde{t}_j)})^\perp = L_j \oplus L_j^\perp$$

with L_j and $L_j^\perp$ infinite-dimensional. We define an operator $\Pi_j: H \to H$ with $\operatorname{spec}_{\mathrm{ess}} = \{-1, 1\}$ by

$$\Pi_{j|\operatorname{range} P_{t_j}^{(\tilde{t}_j)}} = 0\,, \quad \Pi_{j|L_j} = \mathrm{Id}\,, \quad \text{and} \quad \Pi_{j|L_j^\perp} = -\mathrm{Id}\,.$$

Step 4: Finally we define the jump curve

$$A_s^{(j)} := A_s P_s^{(\tilde{t}_j)} + (O_s^j)^* \Pi_j O_s^j \qquad \text{for} \quad t_j \leq s \leq t_{j+1}\,, \tag{4.8}$$

which satisfies the properties 1 and 2. Here the orthogonal projections are chosen in such a way that $P_s^{(\tilde{t}_j)} = O_s^j P_{t_j}^{(\tilde{t}_j)} (O_s^j)^*$ for $t_j \leq s \leq t_{j+1}$. □

Since the dimensions of the kernels do not jump at the discontinuities of the curve, and since the strata $\hat{\mathcal{F}}_*^{(k)}$ are connected, we can insert continuous curve pieces at the discontinuities without changing the spectral flow. This yields a continuous curve $t \mapsto \tilde{A}_t \in \hat{\mathcal{F}}_*$ with the following property:

COROLLARY 4.14. *If each operator $A_t = A_D + C_t$ has an infinite number of eigenvalues on both sides of the real line, then the curve $\{\tilde{A}_t\}$ and the curve $\{\mathcal{R}(A_t)\}$ are homotopic in $\hat{\mathcal{F}}_*$ in the sense that the endpoints are kept in two fixed strata.*

NOTE. The preceding construction leads to a continuous curve $\{\tilde{A}_t\}$ in $\hat{\mathcal{F}}_*$ for any continuous curve of self-adjoint Fredholm operators of the form $A_t = A_D + C_t$. The transformation $\mathcal{R}$, however, leads to a curve in $\hat{\mathcal{F}}_*$ only if the operators of the original curve have an infinite number of eigenvalues on both sides of the real line. Curves of positive or negative semi-bounded operators are mapped by $\mathcal{R}$ in $\hat{\mathcal{F}}_+$, $\hat{\mathcal{F}}_-$, but, by the preceding construction, invariably also in $\hat{\mathcal{F}}_*$.

5. The spectral flow formula.

In this section we shall prove our main result, the equality of the spectral flow and the Maslov index:

THEOREM 5.1 [Spectral flow formula]. *Let A be a closed symmetric operator in a real Hilbert space H with domain D_m and let $\{C_t\}_{t \in I}$ be a continuous family of bounded self-adjoint operators on H. We assume that*

1. *the operator A has a self-adjoint extension A_D with compact resolvent;*
2. *there exists a positive constant a such that*

$$D_m \cap \ker(A^* + C_t - s) = \{0\}$$

for any s with $|s| < a$ and any $t \in [0, 1]$.

Then we have

(5.1) $$\mathrm{sf}(\{A_D + C_t\}) = \mu(\{\gamma(\ker(A^* + C_t)\}, \gamma(D)),$$

where γ denotes the projection of the domain D_M of A^* onto the symplectic space $\beta = D_M/D_m$ of abstract boundary values.

We notice that conditions 1 and 2 are naturally satisfied for operators of Dirac type (i.e. for first-order differential operators with principal symbol of A^2 defining the Riemannian metric). This is valid both over a closed manifold and over a manifold with boundary subject to global elliptic boundary conditions. Clearly, a perturbation which adds a real multiple of the identity will preserve the Dirac type and hence the non-existence of inner solutions ('unique continuation property'). This might, however, not be true for general first-order elliptic differential operators.

We recall that the left side of Formula (5.1) was defined in Definition 4.11. For the right side of Formula (5.1) we recall that $\{\gamma(\ker(A^* + C_t))\}_{t \in I}$ is a continuous family of Lagrangian subspaces of β by the assumptions made and by Theorem 3.8. The pair $(\gamma(\ker(A^* + C_t)), \gamma(D))$ is a Fredholm pair by Proposition 3.5. Hence we have a continuous curve $t \mapsto \Lambda(t) \in \mathcal{FL}_{\Lambda_0}$ in the Fredholm-Lagrangian Grassmannian with $\Lambda(t) := \gamma(\ker(A^* + C_t))$ and $\Lambda_0 := \gamma(D)$. Finally, by Definition 1.4, we have a family of unitary operators $\{W_t : \beta \to \beta\}$ which defines the Maslov index of the curve (see Definition 1.5). The space β is considered as a complex Hilbert space by the almost complex structure which is defined by the form ω introduced in (3.4).

PROOF. We prove the theorem in three steps. First we construct a suitably fine horizontal spacing $\{0 = t_0 < t_1 < \cdots < t_N = 1\}$ and a vertical spacing $\{a_1, \cdots, a_N\}$. Then we show

(5.2) $$\mathrm{sf}(\{\mathscr{R}(A_D + C_{t_{i+1}}) - s\}_{0 \le s \le a_{i+1}/\sqrt{1 + a_{i+1}^2}})$$
$$= \mu(\{\gamma(\ker(A^* + C_{t_{i+1}} - s)\}_{0 \le s \le a_{i+1}}, \gamma(D))$$

for that spacing. This is the main part of the proof. It consists, so to speak, of establishing the coincidence of the spectral flow and the Maslov index for segments of analytic families. This will be done by explicit calculation which identifies the two invariants with the integer $-\sum_{0 \le s \le a_{i+1}} \dim \ker(A_D + C_{t_{i+1}} - s)$. Finally, we show how the general case follows from the special case.

First step: Let t be in $[0, 1]$. We shall construct a suitable horizontal and vertical spacing. We denote the smallest positve eigenvalue of $A_D + C_t$ by $\lambda_1(t)$ and the largest negative one by $\mu_1(t)$. We distinguish two cases:

(I) $\ker(A_D + C_t) = \{0\}$. For the vertical spacing we choose a positive

$$b(t) < \min\{\lambda_1(t), |\mu_1(t)|, a\}.$$

For the horizontal spacing we take a $\delta(t) > 0$ such that the small box is kept free of eigenvalues (see Figure 5), namely

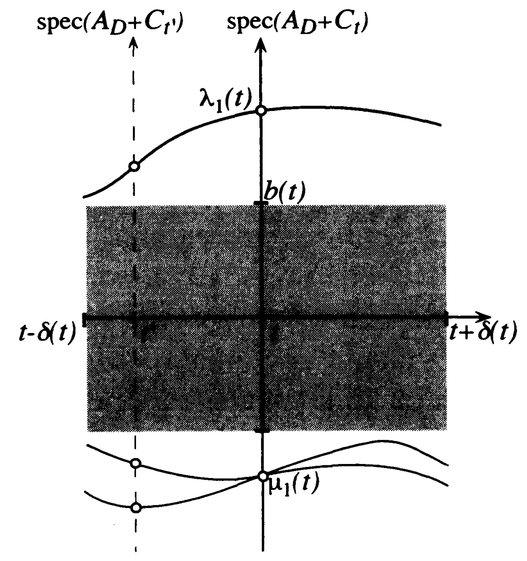


FIGURE 5. Vertical and horizontal spacing for $\ker(A_D+C_t)=\{0\}$

$$\operatorname{spec}(A_D+C_{t'})\cap(-b(t),b(t))=\varnothing \qquad \text{for} \quad t'\in(t-\delta(t),t+\delta(t)).$$

(II) $\ker(A_D+C_t)\neq\{0\}$. In this case we choose a positive

$$b(t)<\min\{\lambda_1(t)/3, |\mu_1(t)|/3, a\}.$$

We choose a $\delta(t)>0$ such that the eigenvalues in a small box are confined by two strips (see Figure 6a), namely

$$\operatorname{spec}(A_D+C_{t'})\cap(b(t),2b(t))=\varnothing,$$

$$\operatorname{spec}(A_D+C_{t'})\cap(-2b(t),-b(t))=\varnothing$$

for $t'\in(t-\delta(t),t+\delta(t))$. This $\delta(t)$ provides the horizontal spacing.

Second step: In the first case, the regular case, we have for each $\tilde{\delta}\leq\delta(t)$

$$\operatorname{sf}(\{\mathscr{R}(A_D+C_{t'})\}_{t-\tilde{\delta}\leq t'\leq t+\tilde{\delta}})=0,$$

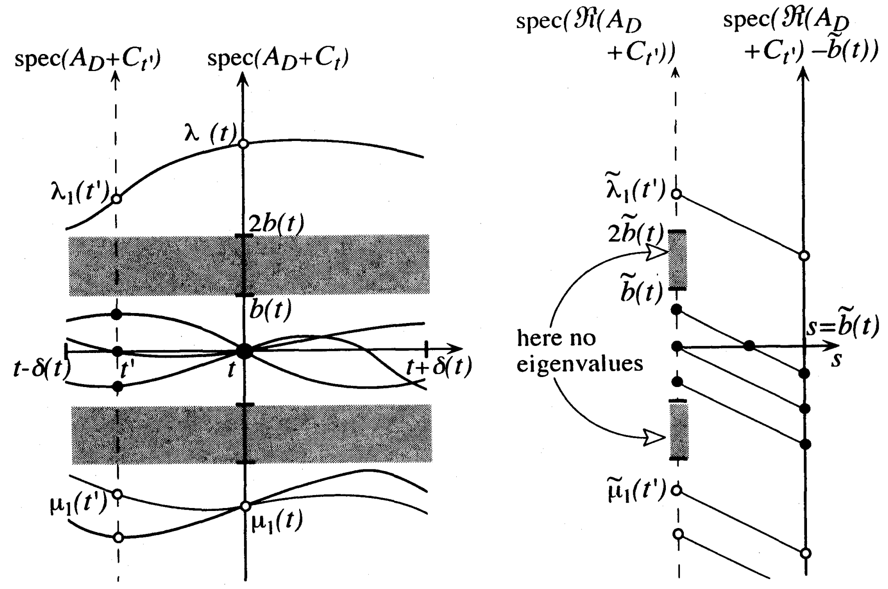
since $A_D+C_{t'}$ is invertible for $t-\delta(t)\leq t'\leq t+\delta(t)$. We also have

$$\mu(\{\gamma(\ker(A^*+C_{t'})\}_{t-\tilde{\delta}\leq t'\leq t+\tilde{\delta}},\gamma(D))=0,$$

since $\gamma(\ker(A^*+C_{t'})\cap\gamma(D)=\{0\}$ for $t-\delta(t)\leq t'\leq t+\delta(t)$.

For the second, singular case, we recall that the eigenvalues of the operator $\mathscr{R}(A_D+C_{t'})-s$ are of the form $\lambda/\sqrt{1+\lambda^2}-s$, where λ is an eigenvalue of the operator $A_D+C_{t'}$. Hence the spectral flow of the family $\{\mathscr{R}(A_D+C_{t'})-s\}_{0\leq s\leq b(t)/\sqrt{1+b(t)^2}}$ equals

$$-\sum_{0\leq s\leq b(t)/\sqrt{1+b(t)^2}}\dim\ker(\mathscr{R}(A_D+C_{t'})-s)=-\sum_{0\leq s\leq b(t)}\dim\ker(A_D+C_{t'}-s)$$



a) Vertical and horizontal spacing for $\ker(A_D+C_t)\neq\{0\}$

b) The spectral flow of the linear family $\{\mathscr{R}(A_D+C_{t'})-s\}_{s\in[0,\tilde{b}(t)]}$, $\tilde{b}(t)=b(t)(1+b(t)^2)^{-1/2}$

FIGURE 6

(see also Figure 6b). We shall show that this integer equals the Maslov index of the family $\{\Lambda(s,t'):=\gamma(\ker(A^*+C_{t'}-s))\}_{0\leq s\leq b(t)}\}$ at the Lagrangian $\Lambda_0:=\gamma(D)$ for each fixed $t'\in(t-\delta(t),t+\delta(t))$. We have $\Lambda(s,t')\cap\Lambda_0=\{0\}$, if $s\notin\operatorname{spec}(A_D+C_{t'})\cap[0,b(t)]$. By our assumption 1, the set $\operatorname{spec}(A_D+C_{t'})\cap[0,b(t)]$ contains only a finite number of elements. Consequently, the intersection of the curve $\{\Lambda(s,t')\}_{0\leq s\leq b(t)}$ with $\gamma(D)$ is nontrivial only at finitely many points; and these points are the eigenvalues λ of the operator $A_D+C_{t'}$ with $0\leq\lambda\leq b(t)$. Clearly, the family $\{\Lambda(s,t')\}_{0\leq s\leq b(t)}$ is a smooth curve.

We determine the quadratic form $Q_{(\Lambda(\lambda,t'),\dot{\Lambda}(\lambda,t'))}$ for all such eigenvalues λ and fixed t':

$$Q_{(\Lambda(\lambda,t'),\dot{\Lambda}(\lambda,t'))}([x],[x]):=\frac{d}{d\theta}\,\omega([x],B_\theta[x])_{|\theta=0}\qquad\text{for}\quad [x]=\gamma(x)\in\beta\,,$$

where $\beta:=D_M/D_m$ denotes our symplectic space of boundary values and $B_\theta\colon\Lambda(\lambda,t')\to\Lambda(\lambda,t')^\perp$ is chosen in such a way that $\{[x]+B_\theta[x]\mid[x]\in\Lambda(\lambda,t')\}=\Lambda(\lambda+\theta,t')$ for θ close to 0. It follows that $B_0=0$.

Let $x\in\ker(A_D+C_{t'}-\lambda)$ or equivalently $\gamma(x)\in\Lambda(\lambda,t')\cap\Lambda_0$, and θ sufficiently small. Then we can choose a smooth family

$$\{u_\theta\in\ker(A^*+C_{t'}-\lambda-\theta)\}$$

such that

$$\gamma(x)+B_\theta(\gamma(x))=\gamma(u_\theta)\quad\text{and}\quad u_0=x\,.$$

Hence

$$\begin{aligned}(5.3)\qquad \omega(\gamma(x), B_\theta(\gamma(x))) &= \langle A^*x, u_\theta - x\rangle - \langle x, A^*(u_\theta - x)\rangle \\ &= \langle (A^* + C_{t'} - \lambda)x, u_\theta - x\rangle - \langle x, (A^* + C_{t'} - \lambda)(u_\theta - x)\rangle = -\langle x, \theta u_\theta\rangle .\end{aligned}$$

Differentiating yields

$$\frac{d}{d\theta} - \langle x, \theta u_\theta\rangle_{|\theta=0} = -\langle x, u_0\rangle = -\langle x, x\rangle < 0$$

for $x \neq 0$; hence $Q_{(A(\lambda,t'),\dot{A}(\lambda,t'))\,|\,A(\lambda,t')\cap\gamma(D)}$ is negative definite.

This implies that the crossings are all regular at $s = \lambda$, the eigenvalue of $A_D + C_{t'}$, so that we can apply Theorem 2.1 and determine the Maslov index by adding the signatures of the crossing forms; and it implies that this signature is just the dimension of the kernel of $A_D + C_{t'} - \lambda$. Hence

$$\begin{aligned}(5.4)\qquad \mu(\{A(s,t')\}_{0\le s\le b(t)}, \gamma(D)) &= \sum_{0\le s\le b(t)} \operatorname{sign} Q_{(A(s,t'),\dot{A}(s,t'))\,|\,A(s,t')\cap A_0} \\ &= -\sum_{0\le s\le b(t)} \dim A(s,t')\cap\gamma(D) = -\sum_{0\le s\le b(t)} \dim\ker(A_D + C_{t'} - s) .\end{aligned}$$

Based on these considerations we can choose the desired horizontal spacing $\{0 = t_0 < t_1 < \cdots < t_N = 1\}$ and vertical spacing $\{a_1 := b(t_1), \cdots, a_N := b(t_N)\}$.

Third step: Let $[t_i, t_{i+1}]$ be one of these small intervals. We consider the two-parameter family $\{\mathscr{R}(A_D + C_t) - s\}$ with $t \in [t_i, t_{i+1}]$ and $s \in [0, a_{i+1}/\sqrt{1 + a_{i+1}^2}]$. We obtain

$$\begin{aligned}(5.5)\qquad &\mathrm{sf}(\{\mathscr{R}(A_D + C_t)\}_{t_i\le t\le t_{i+1}}) + \mathrm{sf}(\{\mathscr{R}(A_D + C_{t_{i+1}}) - s\}_{0\le s\le a_{i+1}/\sqrt{1+a_{i+1}^2}}) \\ &\quad - \mathrm{sf}(\{\mathscr{R}(A_D + C_{t_i}) - s\}_{0\le s\le a_{i+1}/\sqrt{1+a_{i+1}^2}}) = 0\end{aligned}$$

from Theorem 4.2 and

$$\begin{aligned}(5.6)\qquad &\mu(\{\gamma(\ker(A^* + C_t)\}_{t_i\le t\le t_{i+1}}, \gamma(D)) \\ &\quad + \mu(\{\gamma(\ker(A^* + C_{t_{i+1}} - s)\}_{0\le s\le a_{i+1}}, \gamma(D)) \\ &\quad\quad - \mu(\{\gamma(\ker(A^* + C_{t_i} - s)\}_{0\le s\le a_{i+1}}, \gamma(D)) = 0\end{aligned}$$

from Theorem 1.6. The previous step showed that the spectral flow and the Maslov index coincide for linear families. That yields

$$\mathrm{sf}(\{\mathscr{R}(A_D + C_t)\}_{t_i\le t\le t_{i+1}}) = \mu(\{\gamma(\ker(A^* + C_t)\}_{t_i\le t\le t_{i+1}}, \gamma(D))$$

for each small interval $[t_i, t_{i+1}]$. Then additivity under catenation proves the theorem.

□

References

[1] A. ALONSO and B. SIMON, The Birman-Krein-Vishik theory of self-adjoint extensions of semibounded operators, J. Operator Theory **4** (1980), 251–270.

[2] A. ALONSO and B. SIMON, Addenda to "The Birman-Krein-Vishik theory of self-adjoint extensions of semibounded operators", J. Operator Theory **6** (1981), 407.

[3] V. I. ARNOLD, Characteristic class entering in quantization conditions, Funkcional. Anal. i Priložen. **1** (1967), 1–14 (Russian); English transl.: Functional Anal. Appl. **1** (1967), 1–13; French transl.: V. P. Maslov, *Théorie des Perturbations et Méthodes Asymptotiques*, Dunod, Gauthier-Villars (1972), 341–361 (Complément 1).

[4] M. F. ATIYAH, V. K. PATODI and I. M. SINGER, Spectral asymmetry and Riemannian geometry. III, Math. Proc. Cambridge Philos. Soc. **79** (1976), 71–99.

[5] M. F. ATIYAH and I. M. SINGER, Index theory for skew-adjoint Fredholm operators, Inst. Hautes Études Sci. Publ. Math. **37** (1969), 5–26.

[6] B. BOJARSKI, The abstract linear conjugation problem and Fredholm pairs of subspaces, *In Memoriam I.N. Vekua*, Tbilisi Univ. (1979), 45–60 (Russian).

[7] B. BOOSS-BAVNBEK and K. P. WOJCIECHOWSKI, *Elliptic Boundary Problems for Dirac Operators*, Birkhäuser (1993).

[8] ———, Grassmannian and boundary contribution to the ζ-determinant in even dimension (to appear).

[9] U. BUNKE, On the gluing problem for the η-invariant, J. Differential Geom. **41** (1995), 397–448.

[10] S. E. CAPPELL, R. LEE and E. Y. MILLER, On the Maslov index, Comm. Pure Appl. Math. **47** (1994), 121–186.

[11] ———, Selfadjoint elliptic operators and manifold decompositions Part I: Low eigenmodes and stretching, Comm. Pure Appl. Math. **49** (1996), 825–866. Part II: Spectral flow and Maslov index, Comm. Pure Appl. Math. **49** (1996), 869–909. Part III: Applications to Casson's invariant (announced).

[12] H. O. CORDES and J. P. LABROUSSE, The invariance of the index in the metric of closed operators, J. Math. and Mech. **12** (1963), 693–719.

[13] R. G. DOUGLAS and K. P. WOJCIECHOWSKI, Adiabatic limits of the η-invariants. The odd-dimensional Atiyah-Patodi-Singer problem, Comm. Math. Phys. **142** (1991), 139–168.

[14] A. FLOER, A relative Morse index for the symplectic action, Comm. Pure Appl. Math. **41** (1988), 393–407.

[15] B. FUGLEDE, A note on the class of all (unbounded) self-adjoint operators in separable Hilbert space having compact resolvent, Copenhangen (1995), (multiplied).

[16] K. FURUTANI and N. OTSUKI, Spectral flow and intersection number, J. Math. Kyoto Univ. **33** (1993), 261–283.

[17] T. KATO, *Perturbation Theory for Linear Operators*, Springer (1976).

[18] P. A. KIRK and E. P. KLASSEN, Chern-Simons invariants of 3-manifolds decomposed along tori and the circle bundle over the representation space of T^2, Comm. Math. Phys. **153** (1993a), 521–557.

[19] ———, The spectral flow of the odd signature operator on a manifold with boundary, Preprint, Indiana Univ. (1993b).

[20] ———, Computing spectral flow via cup products, J. Differential Geom. **40** (1994a), 505–562.

[21] ———, Analytic deformations of the spectrum of a family of Dirac operators on an odd-dimensional manifold with boundary, Preprint, Indiana Univ. (1994b).

[22] ———, The spectral flow of the odd signature operator and higher Massey products, Preprint, Indiana Univ. (1994c).

[23] J. Leray, Analyse Lagrangiénne et mécanique quantique: Une structure mathématique apparentée aux développements asymptotiques et à l'indice de Maslov, Sér. Math. Pure Appl., I.R.M.P., Strasbourg (1978); English transl.: MIT Press (1981).
[24] R. Nest, Personal communication (1995).
[25] L. Nicolaescu, The Maslov index, the spectral flow, and splittings of manifolds, C. R. Acad. Sci. Paris Sér. 1 **317** (1993), 515–519.
[26] ———, The Maslov index, the spectral flow, and decomposition of manifolds, Duke Math. J. **80** (1995), 485–533.
[27] ———, Generalized symplectic geometries and the index of families of elliptic problems, Preprint, Univ. Michigan (1996).
[28] N. Otsuki and K. Furutani, Spectral flow and Maslov index arising from Lagrangian intersections, Tokyo J. Math. **14** (1991), 135–150.
[29] R. S. Palais, On the homotopy type of certain groups of operators, Topology **3** (1965), 271–279.
[30] J. Phillips, Self-adjoint Fredholm operators and spectral flow, Canad. Math. Bull. **39** (1996).
[31] J. Robbin and D. Salamon, The Maslov index for paths, Topology **32** (1993), 827–844.
[32] R. C. Swanson, Fredholm intersection theory and elliptic boundary deformation problems, I, J. Differential Equations **28** (1978), 189–201.
[33] T. Yoshida, Floer homology and splittings of manifolds, Ann. of Math. **134** (1991), 277–323.

Present Addresses:
Bernhelm Booss-Bavnbek
Institute for Matematik og Fysik, Roskilde University,
DK-4000 Roskilde, Denmark.
e-mail: booss@mmf.ruc.dk

Kenro Furutani
Department of Mathematics, Science University of Tokyo,
Noda-shi, Chiba-Ken, 278–8510 Japan.
e-mail: furutani@ma.noda.sut.ac.jp