SISSA

Scuola Internazionale Superiore di Studi Avanzati

Mathematics Area - PhD course in Geometry and Mathematical Physics

Spectral properties of the Second Variation of an optimal control problem

Candidate:
Stefano Baranzini

Advisor:
Andrei A. Agrachev

Co-advisors:
Ivan Y. Beschastnyi

Academic Year 2021-22

MA PER SEGUIR VIRTUTE E CANOSCENZA · SISSA

Contents

Abstract

In this thesis we study the spectral properties of the Second Variation of an optimal control problem. In particular we focus on three aspects: the asymptotic distribution of the spectrum on the real line, the change in Morse index of an extremal subject to different boundary conditions and the determinant.

We provide, under some regularity assumptions, an exhaustive Weyl-type law for the eigenvalue of the Second Variation.

We prove a formula for the change of Morse index of an extremal satisfying different sets of boundary conditions. We apply it to get iteration formulas for periodic extremals and discretization formulas that reduce the problem of computing the index to a finite dimensional one. Moreover we present some ideas on how to apply the theory to variational problems on graphs.

Finally we provide a way to compute the determinant of the Second Variation in terms of the fundamental solution of a system of ODEs, proving a generalized Hill-type formula. Application to stability are discussed.

Chapter 1

Introduction

Given a function f, how can we find its *maxima* and *minima*? This is a central question in Calculus since its birth and even before in Fermat and Descartes works in the first half of XVII century. Well, the answer is simple (is it really?) look at the *first* and *second derivatives*! The *leitmotiv* of this work is the Second Variation of an optimal control problem and its (spectral) properties.

But ... what is an *optimal control problem*? Well, this is a bit long to explain. A detailed and formal account is given in Appendix B. Let us try to explain it here more informally, like you would do at a dinner table, when someone dares to ask the much dreaded question *"So, what is it that you do exactly?"*.

The situation we want to study is the following: suppose you have a particle (modelling a car, skies, a robot ...) and you want to move it from initial to final state. You might have restrictions on the way you move, depending on the shape of the object you are trying to control. Take for instance a car and the problem of parallel parking: your wheels do not simply turn! This kind of restrictions on motion are sometimes called *non holonomic*, it means that they cannot be written in terms of the position alone. They explicitly depend on the velocity of your particle too. A *control problem* consists in deciding if, given these restrictions on the movement of your particle, it is actually possible to move it from a given initial position to a given final one.

The word *optimal* comes into play when you add to this picture a way to select the *best* trajectory to connect your initial and final state. This is usually done adding the datum of a *functional* to your problem and requiring that the best trajectories are minima (or maxima) of the latter. In Riemannian geometry one would take the kinetic energy. In classical mechanics one would add to that a potential. One could also take the total time itself as the way to weight trajectories, this would be called a *time optimal* problem.

And the *Second Variation*? Well, as Fermat already discovered in the thirties (of the XVII century, of course), minima and maxima are necessarily critical points. However not all critical points are minima or maxima, sufficient conditions were later formulated in terms of the second derivative. It is well known that, when dealing with real functions on open sets $\Omega \subseteq \mathbb{R}^n$, the second derivative at a point p is represented by a symmetric matrix which can be naturally identified with a quadratic form. Thus, for us, the Second Variation will be a quadratic form on a suitable Banach space. It essentially gives a *quadratic* approximation of the functional we use to select the best trajectory in our *optimal control problem*.

The next word I would pick after *Second Variation of an optimal control problem* to describe this work would be *spectrum*. Every result stated in this thesis deals with either the distribution of the spectrum Σ of the Second Variation on the real line or can be used to compute quantities such as:

$$\lim_{\epsilon\to 0^+} \prod_{|\lambda|>\epsilon, \lambda\in\Sigma} \lambda, \qquad \lim_{\epsilon\to 0^+} \sum_{|\lambda|>\epsilon, \lambda\in\Sigma} (\lambda-1)^k,\ k\in\mathbb{N}.$$

In *optimal control* one is particularly interested in the cardinality of the negative part of the spectrum of the second variation, the so called *index*. The reason is that the index tells you in how many independent ways you can reduce the value of your cost for a given trajectory. Generally speaking, if you wish to find the *optimal* strategy to interact with your system you ought to look for trajectory with 0 index.

I think that these few lines are enough to give a taste of things to come. If you are not satisfied with this explanation yet (even though in my experience most people are at this point), well, here is the rest.

The starting goal of my research project was to extend the results in [7] to compute the determinant of the Second Variation. Here by determinant we mean a suitable limit of the product of the eigenvalues of the Second Variation, as in the previous formula. The first step in this direction was to develop a formalism to write the Second Variation in a fashion similar to the fixed point case. However, in doing so, other questions and ideas sprung up. Thanks to Ivan's interest in variational problems on graphs, quantum graphs and so on, Chapter 4 was born. Trying to understand if one could recover the Second Variation of an *LQ* optimal control problem (see [12][Chapter 16]) just by its spectrum (as in the famous "Can you hear the shap of a drum?" problem) led to Chapter 3 and to a general distribution law for the eigenvalues of the Second Variation. Ultimately I managed to prove an extension of the result in [7], this is the content of Chapter 5.

This first Chapter is meant to serve as an account of the main results achieved during my PhD studies at SISSA under the supervision of Prof. Andrei Agrachev and Ivan Beschastnyi. I will go over them in the next sections and try to give a bit of perspective on some of them. Most of what it is presented here is contained in [18, 15, 17]. I will not follow strictly the chapters order and delay the most technical result to the last section. Considerable effort was put in order to adapt not always matching notations, avoid as much as possible repetitions and realize a manuscript with the most coherent presentation I could achieve. For any doubt and possible inconstancy, however, refer to the original sources.

I chose to confine everything that could be labelled as *background material* to Appendices A and B with the sole exception of the differentiation of the Endpoint map which is recalled in Section 2.1 since it serves as motivation for the rest of the Chapter. I will introduce the objects needed when needed. For a systematic exposition please, refer to the appendices.

1.1 Morse Index and Graphs

This section gives an account of the contents of Chapter 4. The main result there is a formula to compute the difference of Morse index of an *extremal* (loosely speaking a *critical point* of the functional we consider) satisfying different sets of boundary conditions. The motivation behind this formula is to apply second variation techniques to variational problems on graphs, much of the material appearing here can be found in [15].

Notice that when we speak of extremals or critical points we usually mean curves in the cotangent bundle of our state space, see Theorem B.1 for a precise statement of this fact. To understand why this is so, think of a problem of classical mechanics or Riemannian geometry. Critical points of the action functional satisfy Euler-Lagrange equations which are *second order* differential equations. Thus the initial position alone is not enough to parametrize all solutions! By graphs here we mean *metric* graph as in the following definition:

Definition 1.1 (Metric graph)**.** A metric graph is a graph $\mathcal{G} = (\mathcal{G}_0, \mathcal{G}_1)$, where $\mathcal{G}_0$ is the set of vertices and $\mathcal{G}_1$ is the set of parametrized edges. Each edge $e \in \mathcal{G}_1$ is parametrized either by a finite interval $[0, l_e]$ for some $l_e > 0$ or by $[0, +\infty)$.

In particular, metric graphs are oriented, i.e. any edge is endowed with an orientation and we can speak of initial (*source*) and final (*target*) point of an edge. Define the maps $s, t : \mathcal{G}_1 \to \mathcal{G}_0$, for each edge $s(e)$ is the initial point whereas $t(e)$ the final one.

We can partition the set of vertices in two using the map just introduced, we define the following:

$$\mathcal{G}_0^f = \{v : \exists! e \in \mathcal{G}, t(e) = v \text{ or } s(e) = v\}, \quad \mathcal{G}_0^c = \mathcal{G}_0 \setminus \mathcal{G}_0^f$$

In plain words $\mathcal{G}_0^f$ is the set of vertices with just one edge either issuing from or reaching it.

Example 1.1. Assume that $\mathcal{G}$ is a compact metric graph and consider the space $H^1(\mathcal{G})$, i.e. continuous functions u with the propriety that $u_e := u|_e \in H^1([0, l_e])$. We impose certain boundary conditions (Dirichlet, Neumann, ...) on $\mathcal{G}_0^f$ and consider the functional:

$$\mathcal{J}(u) = \sum_{e \in \mathcal{G}_1} \int_0^{l_e} \frac{\dot{u}_e^2}{2} - \frac{|u_e|^p}{p} dt, \quad p \in (2, 6). \tag{1.1}$$

We wish to compute the Morse index of the critical points of this functional. To do this we cut the graph and consider the edges separately as a curve γ (after re-parametrization) from $[0, 1]$ to $\mathbb{R}^{\#\mathcal{G}_1}$. The continuity conditions and the extra constraints on $\mathcal{G}_0^f$ can be expressed as boundary conditions for γ, i.e. in the form $(\gamma(0), \gamma(1)) \in N \subseteq \mathbb{R}^{2\#\mathcal{G}_1}$. For example, if we impose homogeneous Dirichlet boundary conditions on $\mathcal{G}_0^f$, N reads:

$$N = \{(q, q') \in \mathbb{R}^{2\#\mathcal{G}_1} : q_e = q_{e'} \text{ if } s(e) = s(e'), q'_e = q'_{e'} \text{ if } t(e) = t(e')\} \cap N_f,$$
$$N_f = \{(q, q') \in \mathbb{R}^{2\#\mathcal{G}_1} : q_e = 0 \text{ if } s(e) \in \mathcal{G}_0^f, q'_e = 0 \text{ if } e \in \mathcal{G}_0^f\}.$$

The choice of the functional in eq. (1.1) is not random. There is an extensive literature concerning the existence and non-existence of ground state (i.e. global minimizers) of the energy in eq. (1.1). The real issue here is that one wishes to consider also non compact graph. For example with a finite number of edges some of which with infinite length, see [3, 2, 4, 23] and [1, 21]. Or *periodic* graphs, i.e. graph with a relatively simple structure repeated infinitely many times, see for example [28] or [29]. Our hope is to be able to extract some necessary optimality conditions, using our techniques, to characterize local minimizers of this problem. Maybe involving finer properties such as number of oscillations or monotonicity. This is currently under investigation jointly with Ivan Beschastnyi.

We can slightly generalize the previous example and consider the following situation: to each edge $e \in \mathcal{G}_1$ of the graph we associate a dynamical system $f_u^e(q)$ on some manifold M^e

and a Lagrangian $\varphi^e(t,u,q)$. We consider boundary conditions $N \subseteq \prod_{e\in\mathcal{G}_1} M^e \times \prod_{e\in\mathcal{G}_1} M^e$ and look for the critical points of:

$$\mathcal{J}(u) = \sum_{e\in\mathcal{G}_1} \int_0^{l_e} \varphi^e(t,u,q_u(t))dt.$$

A typical situation is when all the M^e are the same manifold and we impose a continuity condition at the vertices in $\mathcal{G}_0^c$ and additional ones (Dirichlet, Neumann...) to the vertices in $\mathcal{G}_0^f$.

So, a first step to have effective ways of computing the Morse index for this class of problems is to find effective formulas to compute the Morse index of the problem in eq. (B.2) for general boundary conditions.

The topic is quite classical and several extension of the classical Morse theorem for conjugate points are available. See for example [50]. In the case of fixed endpoints the celebrated Morse theorem states that the index of the second variation is equal to the number of conjugate points along the extremal counted with multiplicity. Let us explain this fact briefly. In some (regular) situation is possible to formulate the condition of being an extremal in terms of an autonomous Hamiltonian system on the cotangent bundle of the state space. Second order conditions are stated in terms of the linearisation (i.e. the differential of the Hamiltonian flow) along the fixed extremal. Loosely speaking these maps give *infinitesimal variations* of the trajectory under consideration. In formulae Morse Theorem reads:

$$ind^-(Q) = \sum_{t\in(0,1)} \dim(\Pi \cap \Phi_t(\Pi)).$$

Here Π denotes the vertical subspace and Φ_t the fundamental solution of the linearisation of the (extremal) flow mentioned above (see Appendix B).

When dealing with more general boundary conditions the Morse Index cannot be express just in terms of conjugates point and correction terms appear. A huge number of works in literature deals with this problem and several intersection indices in the Lagrange Grassmannian have been introduced, see for instance [30, 49, 53, 25].

Here we are not going to pursue this direction. The strategy will be to single out the contribution given by the boundary conditions. We will denote by Q_Π the second variation for the problem with fixed endpoints and Q_N the second variation for boundary conditions given by N. We are going to prove a formula that relates the difference $ind^-(Q_N) - ind^-(Q_\Pi)$ to Maslov index as defined in Definition A.3 and eq. (A.3). More generally we can consider the situation in which an extremal $\lambda_t : [0,1] \to T^*M$ satisfy multiple sets of boundary conditions and try to compute the change of index. For instance, if you think to the example of the graph above, such a formula would provide a way to compute the change of index if we decide to impose an extra Dirichlet condition on a vertex $v \in \mathcal{G}_0^c$ thus shrinking the space of variations.

We can also apply this ideas in a more abstract way to recover iteration formula for the concatenation of periodic extremal and discretization formulas, this is done in Sections 4.1 to 4.3.

We are going now to state the result. First of all let us introduce the *Jacobi equation*. It is a time dependent Hamiltonian system on $T_{\lambda_0}T^*M$ given as follows:

$$\dot{\Phi}_t = Z_t Z_t^* J \Phi_t \tag{1.2}$$

Here J is a coordinate representation of the symplectic form on $T_{\lambda_0}T^*M$ and Z_t is the matrix valued function appearing in the expression of the second variation given in definition 2.1 (see Appendix B and eq. (B.11)). Equation (1.2) can be a bit mysterious: let us shed some light on it. One should think of it as a substitute of the linearisation of the extremal flow appearing in Morse theorem. The reason why we need such system is to cover degenerate situation in which the extremal flow is non-smooth and we cannot really speak of *linearisation* or when we consider *abnormal* extremal (see Appendix B and Definition B.2). Notice, moreover, that the Hamiltonian defining Jacobi equation is non negative. This is because we work with the so called optical Hamiltonian systems as introduced in [16]. These are systems that have a non negativity condition on a Lagrangian subspace: the way in which we compute the the Second Variation allows to separate this positive part from all the rest. What remains is exactly Equation (1.2).

The second object that will appear in the formulas is *Maslov* index of a triple of Lagrange subspaces. Its appearance is ubiquitous in the field and in problems of intersection theory in the Grassmannians (see Appendix A.2). The precise definition is given in eq. (A.3). What it is important to know for now is that it is an invariant of triples of Lagrangian subspaces and measures their relative position assigning to each triple L_0,L_1,L_2 an integer $i(L_0, L_1, L_2)$ satisfying:

$$-\dim(M) \leq i(L_0, L_1, L_2) \leq \dim(M).$$

The third and last definition is that of annihilator of the boundary conditions. It is a submanifold of T^*M, denoted by $Ann(N)$, which plays the same role as the normal bundle in Riemannian geometry. It projects onto N and $Ann(N) \cap T_n^*M$ consists of:

$$Ann(N) \cap T_n^*M = Ann(T_nN) := \{\lambda \in T_n^*M : \lambda(X) = 0, \forall X \in T_nN\}.$$

It turns out that $Ann(N)$ is Lagrangian submanifold of T^*M with the standard symplectic form. This creates a small sign issue when dealing on $T^*(M \times M)$ and $\sigma \oplus (-\sigma)$, as we always do. Finally let us define $A(N)_{\lambda_0}$ as the tangent space of $Ann(N)$ at λ_0. Note that we will drop the subscript when non strictly necessary. See Appendix A and eq. (A.10) for further details.

Theorem 1.1. *Suppose that λ_t is a strictly normal extremal (see Definition B.2 and the beginning of next Section) satisfying the transversality conditions for two distinct submanifolds N and $\tilde{N}$ (see Theorem B.1). Denote by Φ_t the fundamental solution of eq. (1.2) and by $\Gamma(\tilde{\Phi}_*\Phi_t)$ the graph of its composition with the differential $\tilde{\Phi}_*$ of the re-parametrization flow given in eq. (B.11). Let Q_N and $Q_{\tilde{N}}$ be the respective second variation, then it holds:*

$$ind^- Q_{\tilde{N}} - ind Q_N = i(A(N), \Gamma(\tilde{\Phi}_*\Phi_1), A(\tilde{N})) + d_0. \tag{1.3}$$

The number d_0 appearing above is determined by the following relation:

$$d_0 = \dim\left(A(N) \cap \Gamma(\tilde{\Phi}_*\Phi_1) \,/\, A(N) \cap \Gamma(\tilde{\Phi}_*\Phi_1) \cap A(\tilde{N})\right) - \dim\left(TN \,/\, TN \cap T\tilde{N}\right)$$

Remark 1.1. We always work with the symplectic form $(-\sigma) \oplus \sigma$ when dealing with graphs of symplectomorphisms. In particular the Maslov index $i(\cdot, \cdot, \cdot)$ appearing in eq. (1.3) must be computed with said form.

Remark 1.2. If we assume that the maximum condition of PMP defines a C^2 function H^t in the neighbourhood of the extremal λ_t, the symplectomorphism $\tilde{\Phi}_*\Phi_1$ is the fundamental solution of the linearisation of the flow generated by $\vec{H}^t$ along λ_t. See for example [12][Proposition 21.3].

1.2 Functional determinants and Hill-type formulae

In this section we present the topic of Chapter 5. We will stick to a fairly regular situation. First of all we will assume that the extremal we are fixing satisfies *Legendre* strong condition, as in the previous section. Together with this *extremal*, always comes a control function $\tilde{u}$. It represents the optimal strategy we have devised to stir the particle from the initial to the final position. We will assume that $\tilde{u}$ is a regular point of the Endpoint map. That is, λ_t is a strictly normal extremal as in definition B.2. All this notions are explained in Appendix B. Notice that, the condition of being a regular point of the Endpoint map, is automatically satisfied by extremals of mechanical, Riemannian, and LQ problems. So there is no loss of generality in thinking of those for the moment. The already mentioned *Legendre* condition is a strict convexity condition, with respect to the velocity, of the maximized Hamiltonian along the extremal. It satisfied, for instance, by the already mentioned class of problems.

Under these hypothesis the Second Variation can be naturally interpreted as the Hessian of the cost $\mathcal{J}$ we are considering, restricted to the smooth manifold given by a level set of the Endpoint map.

Our goal is to compute the *determinant* of the Second Variation. We will prove a formula which relates this determinant to the fundamental solutions of an ODE system in a finite dimensional space, much in the spirit of Gelfand-Yaglom Theorem.

Under the hypothesis above, the Second Variation Q is induced by a Fredholm operator and can be written, after appropriate choice of scalar product, as $1+K$ where K is a compact operator. Various ways of defining a *determinant* function on spaces of operators of this form can be found in the literature, going all the way back to the works of Poincaré, Fredholm and Hilbert. If the operator K one considers is in the so called *trace class*, i.e. the sequence of its eigenvalues (with multiplicity) gives an absolutely convergent series, a definition of determinant which involves an infinite product of its eigenvalues is possible.

In our case, however, the classical approach is not immediately applicable since, generically, the operators one encounters when considering Second Variations are not trace class, as it will be clear from Theorem 1.5. We will see that, under some technical assumptions, this class of operators either has a particularly symmetric spectrum or it is trace class. This symmetry allows us to talk about trace and determinant of the operators K and $1+K$ in the sense of *principal limits*. A similar approach has been independently adopted in the works [38, 37] to study the spectrum of Hamiltonian systems. The difference here is the choice of projectors: we use directly the eigenprojectors of the Second Variation.

There are other ways to define the determinant of the Second Variation, of course. For instance, the theory of regularized determinants, see [55], provides a way to define regularized determinants for all Fredholm operators of the form $1+K$, enforcing some condition on the summability of the singular values of K. One introduces the so-called *trace ideals* $\mathfrak{S}_p$ for $p \geq 1$ and defines an appropriate analytic function on each of them which mimics the properties of the determinant. Trace ideals are defined as follows: let K be a compact operator and $\{s_n(K)\}_{n\in\mathbb{N}}$ be the collection of its (ordered) singular values, K belongs to $\mathfrak{S}_p$ if $\sum_{n\in\mathbb{N}} s_n^p(K) < \infty$. For every $m \in \mathbb{N}$ one maps the class $\mathfrak{S}_n$ to $\mathfrak{S}_1$, i.e. trace class operators, and consider the (*analytic*) function:

$$\mathfrak{S}_m \ni K \mapsto \det_1\left((1+K)\exp\left(\sum_{j=1}^{m-1}\frac{(-K)^j}{j}\right)\right) := \det_m(1+K)$$

This definitions preserves many of the properties of finite dimensional determinants, however introduces a heavy non linearity! Moreover, notice that the ideals $\mathfrak{S}_n$, $n \in \mathbb{N}$, are increasing but the various definition of determinants do not agree. If $K \in \mathfrak{S}_m$ with $m < n$ then $\det_n(1+K) \neq \det_m(1+K)$.

Our approach, which is of course less general, however, provides an actual extension of the definition of determinant given for trace class operators. It involves just the infinite product of the eigenvalues and thus, whenever the compact part of the second variation is trace class, it gives exactly the usual determinant $\det_1$ defined for trace class operators.

It is worth pointing out two main features of our construction. First of all it relates the determinant of $1+K$ to the fundamental solution of a finite dimensional system of ODEs. This provides a way to actually compute the determinant and allows to recover some classical results such as Hill's formula for periodic trajectories. This kind of formulas have important applications since allow to relate variational properties of an extremal (i.e. the eigenvalues of the second variation) to dynamical properties such as *stability*, which are determined by the eigenvalues of the linearisation of the Hamiltonian system of which the extremal we are considering is solution. For details about Hill formula and stability issues we just mention the papers [20] and [40] and refer to the discussion at the beginning of Chapter 5.

We stress that our results are formulated in a quite general framework which encompasses Riemannian, sub-Riemannian and Finsler geometry, mechanical systems on manifolds to name a few. Moreover the techniques can be applied without virtually any modifications to treat constrained variational problems on graphs as done in Chapter 4 to compute the Morse index.

Yet another possible definition of determinant, this time for (pseudo-)differential operators, is the following. To a self-adjoint differential operator L with a set of boundary condition B one associates a $\zeta-$ function defined as:

$$\zeta(s) = \sum_{\lambda \in \mathrm{spec}(L)} \lambda^{-s}, \ s \in \mathbb{C} \text{ with } \Re(s) \text{ sufficiently big.}$$

The series converges on some half-plane $\Re(s) > c_0$ and is then extended to a unique meromorphic function on the whole plane by analytic continuation. The determinant of L is then defined to be $-\log(e^{\zeta'(s)})|_{s=0}$. The reader is referred to [31] for details. Let us just point out that $\zeta-$regularization techniques are the ones employed to obtain expression for quantities such as the *ratio* of the determinants of two differential operators. These constructions are carried out in many paper such as [46] for Sturm-Liouville problems and [32] in the general framework of elliptic operators on section of vector bundles. A relation between regularized determinants and zeta-regularization is given in [34]. In principle one could recover similar results to the ones obtained in [33] for the determinant of elliptic operators on quantum graphs or [45] for Sturm-Liouville differential operators. However our formulas appearance would be slightly different since we make a different choice of reference operator.

We can interpret the expression of the second variation coming from Chronological Calculus (see [12][Chapter 3]) as the composition of a differential operator and the inverse of another one, canonically associated to the problem. As reference we use some symplectic version of the parallel transport operator along our extremal. A precise definition of what we mean by *parallel transport* will be given later on, in Appendix A and Section 5.1, where all the objects mentioned above will be properly defined.

Let us state now our result. Denote by $m(\lambda)$ the multiplicity of an eigenvalue λ of Q, we define the determinant of Q as:

$$\det(Q) = \lim_{\epsilon\to 0} \prod_{|\lambda-1|>\epsilon} \lambda^{m(\lambda)}, \text{ where } \lambda \in \operatorname{Spec}(Q).$$

We introduce now a variation of Jacobi equation as given in eq. (1.2). To define it we use the two matrix valued functions $Z_t \in \mathrm{Mat}_{k\times 2\dim(M)}(\mathbb{R})$ and $H_t \in \mathrm{Mat}_{k\times k}(\mathbb{R})$ introduced before. The precise way to compute them is explained in eq. (B.11). Let δ^s be dilations of a parameter $s \in \mathbb{R}$ of the fibre, i.e. the kernel of the differential of the natural projection $\pi : T^*_{\lambda_0}M \to M$. The dilation δ^s is defined by the relations $\pi_*(\delta^s w) = \pi_*(w)$ for all $w \in T_{\lambda_0}T^*M$ and $\delta^s v = sv$ for all $v \in \ker(\pi_*)$. Let J be some representation in coordinates of the standard symplectic form on $T_{\lambda_0}T^*M$, using the maps just introduce let us define the following quadratic form:

$$\beta_s(\lambda) = \frac{1}{2}\langle \lambda, J\delta^s Z_t H_t^{-1}(\delta^s Z_t)^* J\lambda\rangle, \quad \lambda \in T_{\lambda_0}T^*M.$$

We will call *Jacobi equation* the following ODEs system on $T_{\lambda_0}T^*M$:

$$\dot\lambda = \vec\beta_s(\lambda), \quad \lambda \in T_{\lambda_0}T^*M.$$

and denote its fundamental solution at time t as Φ^s_t. Notice that for $s = 1$ we recover exactly eq. (1.2).

The last maps we will need is a family of symplectomorphism of $T_{\lambda_0}T^*M \times T_{\lambda_1}T^*M$ which depend on the choice of a scalar product on each space. Let g_0 and g_1 be two scalar products on $T_{\lambda_0}T^*M$ and $T_{\lambda_1}T^*M$ respectively. Assume that at each λ_i, $\Pi_i := \ker(\pi_*)$ has a Lagrangian orthogonal complement wrt g_i which we denote by $\Pi_i^\perp$. Denote by pr_V the orthogonal projection onto V and set:

$$\begin{aligned}
A^s_0(\eta) &= \eta + (1-s)J_0^{-1} pr_{\Pi_0^\perp}\eta, \quad \eta \in T_{\lambda_0}(T^*M),\\
A^s_1(\eta) &= \eta + (1-s)(J_1^{-1} + \tilde\Phi_* \circ pr_{\Pi_0} \circ \tilde\Phi_*^{-1}) pr_{\Pi_1^\perp}\eta, \quad \eta \in T_{\lambda_1}(T^*M).
\end{aligned}$$

This maps preserve the fibre Π_0 and Π_1 respectively and move any Lagrangian subspace closer and closer to the fibre as s goes to infinity.

The datum of the boundary condition is encoded in a Lagrangian submanifold of $\big(T^*(M\times M), (-\sigma)\oplus\sigma\big)$, the annihilator of N already appearing in the statement of Theorem 1.1. Take a sub-manifold $N \subseteq M\times M$, the explicit definition is as follows:

$$Ann(N) = \bigcup_{q\in N}\{(\lambda_0,\lambda_1) \in T^*_q(M\times M) : \langle\lambda_0, X_0\rangle = \langle\lambda_1, X_1\rangle, \forall (X_0,X_1) \in T_qN\}.$$

In light of PMP (see Theorem B.1), critical points of $\mathcal{J}$ with boundary conditions given by N, lift to the cotangent bundle to curves λ_t such that $(\lambda_0,\lambda_1) \in Ann(N)$. Fix now a complement to $T_{(\lambda_0,\lambda_1)}Ann(N)$, say V_N, and denote by π_N the projection onto this latter subspace with kernel the tangent space to the annihilator.

We are ready now define a function that plays the role of the *characteristic polynomial* of Q. For a map f denote by $\Gamma(f)$ its graph, set:

$$\mathfrak{p}_Q(s) = \det(\pi_N|_{\Gamma(A^s_1\tilde\Phi_*\Phi^s_1 A^s_0)}).$$

With this notation our main result reads as follows:

Theorem 1.2. *Assume that the matrix $-H_t > \alpha$ for some $\alpha > 0$ (*Legendre strong condition*) and that at least one of the following holds:*

i) the maps $t \mapsto Z_t$ and $t \mapsto H_t$ are piecewise analytic in t;

ii) the dimension of the space of controls is $k \leq 2$;

iii) the operator $I - Q$ is trace class;

Let $\lambda \in \mathrm{Spec}(Q)$ and $m(\lambda)$ be its multiplicity, the limits:

$$\det(Q) = \lim_{\epsilon \to 0} \prod_{|\lambda - 1| > \epsilon} \lambda^{m(\lambda)}, \quad \mathrm{tr}(Q-1) = \lim_{\epsilon \to 0} \sum_{|\lambda-1|>\epsilon} m(\lambda)(\lambda - 1).$$

are well defined and finite. Moreover, for almost any choice of metrics g_0 and g_1 we have that $\mathfrak{p}_Q(0) \neq 0$ and that:

$$\det(1 + s(Q-1)) = \mathfrak{p}_Q(0)^{-1} e^{s(\mathrm{tr}(Q-1) - \mathfrak{p}_Q(0)')} \mathfrak{p}_Q(s).$$

Remark 1.3. The hypothesis about the regularity of Z_t and H_t are needed to obtain the asymptotic for the spectrum of $Q - 1$ that guarantee the existence of the trace and of the determinant as limits. They can be weakened somehow by requiring that the skew-symmetric $k \times k$ matrix $Z_t^* J Z_t$ is continuously diagonalizable (see [7]).

Remark 1.4. The constants $\mathfrak{p}_Q(0)$, $\mathfrak{p}_q(0)'$ and $\mathrm{tr}(Q-1)$ are completely explicit and are given in terms of iterated integrals in Lemmas 5.3 and 5.4.

In particular we have the following corollary:

Corollary 1.1. *Under the assumption above, the determinant of the second variation Q satisfies:*

$$\det(Q) = \mathfrak{p}_Q(0)^{-1} e^{\mathrm{tr}(Q-1) - \mathfrak{p}_Q(0)'} \det(\pi_N|_{\Gamma(\Psi_t)}).$$

Where $\Psi_t = \tilde{\Phi}_* \Phi_1$ and coincides with the fundamental solution of the linearisation of the extremal flow, whenever the latter is defined.

1.3 Characterization of the Second Variation

Alas, we turn to the last topic that will be touched in this thesis. In this section we are going to overview the contents of Chapters 2 and 3. The common theme is the Second Variation for Dirichlet boundary conditions. As explained and done in the previous section, we will not work with just a trajectory, but with its lift to the cotangent bundle. This lift is given by PMP (Theorem B.1) and we denote it by λ_t. Let $\tilde{u}$ be the relative optimal control.

In Appendix B is shown that the second variation at a critical point $(\lambda_t, \tilde{u})$ of an optimal control problem is a quadratic form (or a family of quadratic forms when the corank of the critical point is ≥ 2) defined on the kernel of the first differential. In Proposition B.2 the following formulas to represent the derivatives of the endpoint map and the second variation (here we use the notation summarized in eq. (B.11)) are given:

$$d_{\tilde{u}} E^t_{q_0}(v) = \pi_* (\tilde{\Phi}_t)_* \int_0^t Z_\tau v_\tau d\tau$$

$$\lambda_t d^2_{\tilde{u}} E^t_{q_0}(v, v) = \int_0^t \left(\langle H_\tau v_\tau, v_\tau \rangle + \int_0^\tau \sigma_{\lambda_0}(Z_\theta v_\theta, Z_\tau v_\tau) d\theta \right) d\tau.$$

One should not be bother too much by these formulas. A quick glance and plain acceptance is advised for the moment. A rigorous derivation is given in Section 2.1 and appendix B. The message one ought to take home by looking at this equations is that the second variation is expressed in integral form. Z_τ and H_τ are function from the interval $[0,t]$ to the space of $k\times 2\dim(M)$ and $k\times k$ matrices respectively, where $k\leq \dim(M)$ is the number of control parameters of the system. They should be interpreted as *coefficients* of the Second Variation. From now on we will forget about where this quadratic form comes from and how we obtained it and study it just as a quadratic form on $L^2([0,t],\mathbb{R}^k)$.

The main focus of this section is the study of the second term appearing in the expression of the second variation, namely on operators of the form:

$$K(v)(\tau) = \int_0^\tau \sigma_{\lambda_0}(Z_\tau\cdot, Z_\theta v_\theta)d\theta. \tag{1.4}$$

They are always compact. The domain of definition is usually taken to be $L^\infty([0,t],\mathbb{R}^k)$ since one works with Lipschitz continuous curves. However it is clear that they extend to continuous operators on $L^2([0,t],\mathbb{R}^k)$. For this reason we will always work on $L^2([0,t],\mathbb{R}^k)$ with the standard Hilbert structure.

We are going to present here a result about the asymptotic distribution of their *spectrum* and a simple characterization of this class. It turns out that the following properties identify the operators as in eq. (1.4):

i) there exists a finite co-dimensional subspace of $L^2([0,1],\mathbb{R}^k)$, which we call $\mathcal{V}$, on which K becomes a self-adjoint operator, i.e. :

$$\langle u, Kv\rangle = \langle Ku, v\rangle \quad \forall\, u,v\in\mathcal{V}, \tag{1.5}$$

ii) K is an Hilbert-Schmidt operator with an integral kernel of a particular form, namely:

$$K(v)(t) = \int_0^t V(t,\tau)v(\tau)d\tau, \quad v\in L^2([0,1],\mathbb{R}^k). \tag{1.6}$$

Where $V(t,\tau)$ is a matrix whose entries are L^2 functions. We call the class of operator satisfying this last condition *Volterra-type* operators.

This characterization is summarized in the following theorem. Let $\mathcal{A}$ denote the skew-symmetric part of K, which we assume to satisfy eqs. (1.5) and (1.6):

$$\mathcal{A} = \frac{1}{2}\Big(K - K^*\Big).$$

Denote by $\mathrm{Im}(\mathcal{A})$ the image of $\mathcal{A}$.

Theorem 1.3. *Let be K an operator satisfying eq. (1.5) and eq. (1.6). Then $\mathcal{A}$ has finite rank and completely determines K. More precisely, if $\mathcal{A}$ has rank $2m$ and is represented as:*

$$\mathcal{A}(v)(t) := \frac{1}{2}Z_t^*\mathcal{A}_0\int_0^1 Z_\tau v(\tau)dt,$$

for a skew-symmetric $2m\times 2m$ matrix $\mathcal{A}_0$ and a $2m\times k$ matrix Z_t then:

$$K(v)(t) = \int_0^t Z_t^*\mathcal{A}_0 Z_\tau v(\tau)d\tau.$$

We can rephrase the above statement in terms of the second variation at a strictly normal critical point (see Definition B.2). We have the following:

Theorem 1.4. *Suppose $\mathcal{V} \subset L^2([0,1],\mathbb{R}^k)$ is a finite codimension subspace and K and operator satisfying eqs. (1.5) and (1.6). Then $(K,\mathcal{V})$ can be realized as the second variation of an optimal control problem at a strictly normal regular extremal. To any such couple we can associate a triple $((\Sigma,\sigma),\Pi,Z)$ consisting of:*

i) a finite dimensional symplectic space (Σ,σ);

ii) a Lagrangian subspace $\Pi \subset \Sigma$;

iii) a linear map $Z : L^2([0,1],\mathbb{R}^k) \to \Sigma$ such that $\mathrm{Im}(Z)$ is transversal to the subspace Π.

This triple is unique up to the action of $\mathrm{stab}_\Pi(\Sigma,\sigma)$, the group of symplectic transformations that fix Π. Any other triple is given by $((\Sigma,\sigma),\Pi,\Phi\circ Z)$ for $\Phi \in \mathrm{stab}_\Pi(\Sigma,\sigma)$.

Vice versa any triple $((\Sigma,\sigma),\Pi,Z)$ as above determines a couple $(K,\mathcal{V})$. We can define the skew-symmetric part $\mathcal{A}$ of K as:

$$\langle \mathcal{A}u, v\rangle = \sigma(Zu, Zv),\ \forall u,v \in L^2([0,1],\mathbb{R}^k),$$

$\mathcal{A}$ determines the whole operator K and its domain is recovered as $\mathcal{V} = Z^{-1}(\Pi)$.

Before giving the precise statements of the next result about the spectral asymptotic, let us pause a little the discussion to recall some general facts about the spectrum of compact operators. Mainly to fix notation and give a couple of definitions.

Given a compact self-adjoint operator K on an Hilbert space $\mathcal{H}$, we can define a quadratic form setting $\mathcal{Q}(v) = \langle v, K(v)\rangle$. The eigenvalues of $\mathcal{Q}$ are by definition those of K and we will denote $\Sigma_\pm(\mathcal{Q})$ the positive and negative parts of the spectrum of $\mathcal{Q}$.

By the standard spectral theory of compact operators (see [54]) the non zero eigenvalues of K are either finite or accumulate at zero and their multiplicity is finite. Consider the positive part of the spectrum of $\mathcal{Q}$, $\Sigma_+(\mathcal{Q})$ and $\lambda \in \Sigma_+(\mathcal{Q})$. Denote by m_λ the multiplicity of the eigenvalue λ. We can introduce a monotone non increasing sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ indexing the eigenvalues of K, requiring that the cardinality of the set $\{\lambda_n = \lambda\} = m_\lambda$ for every $\lambda \in \Sigma_+(\mathcal{Q})$.

This will be called the monotone arrangement of $\Sigma_+(\mathcal{Q})$. We can perform the same construction indexing by $-n$, $n \in \mathbb{N}$, the negative part of the spectrum $\Sigma_-(\mathcal{Q})$. This time we require that the sequence $\{\lambda_{-n}\}_{n\in\mathbb{N}}$ is non decreasing. Provided that $\Sigma_\pm(\mathcal{Q})$ are both infinite, we obtain a sequence $\{\lambda_n\}_{n\in\mathbb{Z}}$.

The next definition is non-standard and the term *capacity* probably already used in too many different contests. The terminology is inspired by [7] and extended here to cover our case. It will essentially encode the first term of the asymptotic of $\{\lambda_n\}_{n\in\mathbb{Z}}$.

Definition 1.2. Let $\mathcal{Q}$ be a quadratic form $\mathcal{Q}$ on a Hilbert space $\mathcal{H}$ and $j \in \mathbb{N}$

i) if j is odd, $\mathcal{Q}$ has $j-$capacity $\xi > 0$ with reminder of order $\nu > 0$ if $\Sigma_+(\mathcal{Q})$ and $\Sigma_-(\mathcal{Q})$ are both infinite and:

$$\lambda_n = \frac{\xi}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n \to \pm\infty,$$

ii) if j is even, $\mathcal{Q}$ has $j-$capacity (ξ_+,ξ_-) with remainder of order $\nu>0$ if either both $\Sigma_+(\mathcal{Q})$ and $\Sigma_-(\mathcal{Q})$ are infinite and:

$$\lambda_n = \frac{\xi_+}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n\to+\infty,$$
$$\lambda_n = \frac{-\xi_-}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n\to-\infty,$$

where $\xi_\pm \geq 0$ or if at least one between $\Sigma_+(\mathcal{Q})$ and $\Sigma_-(\mathcal{Q})$ is infinite and the relative monotone arrangement satisfies the corresponding asymptotic relation;

iii) if the spectrum is finite or $\lambda_n = O(n^{-\nu})$ as $n\to\pm\infty$ for any $\nu>0$, we say that $\mathcal{Q}$ has $\infty-$capacity.

The behaviour of the sequence $\{\lambda_n\}_{n\in\mathbb{Z}}$ is closely related to the following counting functions:

$$C_j^+(n) = \#\{l\in\mathbb{N} : 0 < \frac{1}{\sqrt[j]{\lambda_l}} < n\} \quad C_j^-(n) = \#\{l\in\mathbb{N} : -n > \frac{-1}{\sqrt[j]{|\lambda_{-l}|}} > 0\}$$

The requirement of Definition 1.2 for the $j-$capacity can be translated into the following asymptotic for the functions $C_j^\pm(n)$:

$$C_j^\pm(n) = \frac{\xi_\pm}{\pi}n + O(n^{1-\nu}) \quad \text{as} \quad n\to\pm\infty$$

We mention here some of the properties of the $j-$capacity. Without loss of generality we state the properties for the positive part of the spectrum, analogue results hold for the negative one.

i) *(Homogeneity)* if $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are quadratic forms on two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of $j-$capacity ξ_1 and ξ_2 respectively with the same remainder ν, then $a\mathcal{Q}_1$ has $j-$capacity $a\xi_1$ and the sum $\mathcal{Q}_1\oplus\mathcal{Q}_2$ on $\mathcal{H}_1\oplus\mathcal{H}_2$ has $j-$capacity $(\sqrt[j]{\xi_1}+\sqrt[j]{\xi_2})^j$ both with remainder ν.

ii) *(Independence of restriction)* If $\mathcal{V}\subseteq\mathcal{H}$ is a subspace of finite codimension then $\mathcal{Q}$ has $j-$capacity ξ with remainder ν if and only if its restriction to $\mathcal{V}$ has $j-$capacity ξ with remainder ν.

iii) *(Additivity)* if $\mathcal{Q}_1$ has $j-$capacity ξ with remainder ν and $\mathcal{Q}_2$ has $j-$capacity 0 with remainder of the same order ν, then their sum $\mathcal{Q}_1+\mathcal{Q}_2$ has the capacity ξ with remainder $\nu' = \frac{(j+\nu)(j+1)}{j+\nu+1}$

We will focus now on quadratic forms $\mathcal{Q}$ coming from operators of the form given in eq. (1.4) in light of Theorem 1.3. After a change of coordinates we can assume without loss of generality that the symplectic form is the standard one and thus, using eq. (A.1), we will consider the following quadratic form on $L^2([0,1],\mathbb{R}^k)$::

$$\mathcal{Q}(v) = \langle v, K(v)\rangle = \int_0^1\int_0^t \langle Z_t v(t), J Z_\tau v(\tau)\rangle d\tau dt. \tag{1.7}$$

We will always make the assumption that Z_t is a $2n\times k$ matrix which depends piecewise analytically on the parameter $t\in[0,1]$ through out this section. Let f be a smooth function on $[0,1]$ and let $k \in \mathbb{N}$, denote by $f^{(k)} = \frac{d^k f}{dt^k}$ the $k-$th derivative with respect to t. For $j \geq 1$ define the following matrix valued functions:

$$A_j(t) = \begin{cases} \big(Z_t^{(k)}\big)^* J Z_t^{(k)} & \text{if } j = 2k+1 \\ \big(Z_t^{(k-1)}\big)^* J Z_t^{(k)} & \text{if } j = 2k \end{cases} \tag{1.8}$$

We use ρ_t to denote any eigenvalue of the matrix $A_j(t)$. If $j = 2k$, define:

$$\mu_{t,2k}^+ := \sum_{\rho_t : \rho_t > 0} \sqrt[2k]{\rho_t} \qquad \mu_{t,2k}^- := \sum_{\rho_t : \rho_t < 0} \sqrt[2k]{|\rho_t|}.$$

For odd indices, A_{2k-1} is skew-symmetric and thus the spectrum is purely imaginary. So we define the function:

$$\mu_{t,2k-1} = \sum_{\rho_t : -i\rho_t > 0} \sqrt[2k-1]{-i\rho_t}.$$

We are now ready to state the following result:

Theorem 1.5. *Let $\mathcal{Q}$ be the quadratic form in eq. (1.7). $\mathcal{Q}$ has either $\infty-$capacity or $j-$capacity with remainder of order $\nu = 1/2$. More precisely, let $j \geq 1$ be the lowest integer such that $A_j(t)$ is not identically zero, then*

i) if $j = 2k-1$, the $(2k-1)-$capacity ξ is given by:

$$\xi = \left(\int_0^1 \mu_{t,2k-1} dt \right)^{2k-1},$$

and thus for $n \in \mathbb{Z}$ sufficiently large:

$$\lambda_n = \frac{\left(\int_0^1 \mu_{t,2k-1} dt \right)^{2k-1}}{(\pi n)^{2k-1}} + O(n^{-2k+1/2}).$$

ii) if $j = 2k$, the $2k-$capacity (ξ_+, ξ_-) is given by:

$$\xi_\pm = \left(\int_0^1 \mu_{t,2k}^\pm dt \right)^{2k},$$

and thus for $n \in \mathbb{Z}$ sufficiently large:

$$\lambda_n = \frac{\left(\int_0^1 \mu_{t,2k}^\pm dt \right)^{2k}}{(\pi n)^{2k}} + O(n^{-2k-1/2}).$$

iii) if $A_j(t) \equiv 0$ for any j then $\mathcal{Q}$ has $\infty-$capacity.

Remark 1.5. It is worth remarking that in Theorem 1 of [7] the order of the remainder for the $1-$capacity was a little better, $2/3$ and not $1/2$.

In particular, combining with Theorem 1.3 and assuming the kernel $V(t,\tau)$ in eq. (1.6) is an analytic function in t and τ, we have the following bound on the $1-$capacity in terms of $\mathcal{A}$, the skew-symmetric part of K.

Corollary 1.2. *Let Σ be the spectrum of $\mathcal{A}$, if the matrix $V(t,\tau)$ in eq. (1.6) is analytic, the $1-$capacity of K can be bound by*

$$\xi \leq 2\sqrt{m} \sqrt{\sum_{\rho \in \Sigma : -i\rho > 0} -\rho^2} \leq 2\sqrt{m} \sum_{\rho \in \Sigma : -i\rho > 0} |\rho|.$$

Chapter 2

The Second Variation

This chapter contains the proofs of theorems Theorem 1.3, Theorem 1.4 and the construction we will employ several times through out the work to reduce an optimal control with moving boundary conditions to a fixed points one. Much of the notation used here is introduced either in Appendix B or Section 1.3, however we will soon recall briefly some facts about the differentiation of the Endpoint map to keep the chapter as self contained as possible. The topics presented here are essentially contained in [18] and [15].

The chapter structure is the following. In Section 2.1 we recall how to differentiate the Endpoint map of a control system for Dirichlet boundary condition. The topic is quite standard and further references can be found in [5, 8, 12]. In Section 2.2 we study the integral representation of the second variation thus obtained and characterize the class of integral operators one can realize as second variation of an optimal control problem. Finally Section 2.3 describe how we reduce the second variation with moving endpoints to the well known case of Dirichlet boundary conditions.

2.1 Second Variation with fixed endpoints

Consider an optimal control problem on a smooth manifold M. Roughly speaking the situation is as follows: we specify a family of Lipschitz curves imposing some restriction on its velocity and we want to find *the best strategy* (with respect to some cost functional) to stir a point q_0 to a point q_1. This problem is formalized as follow (see Appendix B for more details). We take a family of vector fields $f_u(q)$ on M depending on some parameter $u \in U \subseteq \mathbb{R}^k$ and look at the solutions of

$$\dot{q}_u = f_{u(t)}(q_u)$$

for $u \in L^\infty([0,1], U)$. This will be our family of Lipschitz curve. The weight we assign to each trajectory will be determined by a smooth function φ_t as follows:

$$\mathcal{J}(u) = \int_0^1 \varphi_t(u, q_u(t), t)dt$$

We denote by $E^t_{q_0}(u)$ the map that takes a function $u(\cdot)$, the *control*, and gives back the solution of the above equation at time t with initial condition q_0.

Any optimal control problem can be reduced to the study of an appropriate Endpoint map simply by adding the cost as an extra variable or can be seen as a problem of constrained optimization with constraints given by the level set of the Endpoint map (see

[12, 43, 14]). We are going now to recall the integral representation for the derivatives of the Endpoint map, complete details are given in Appendix B.

We consider the following family of functions on T^*M:

$$h_u : T^*M \to \mathbb{R}, \quad h_u(\lambda) = \langle \lambda, f_u \rangle + \nu\varphi(u, \pi(\lambda)), \quad \nu \leq 0.$$

By Pontryagin Maximum Principle (Theorem B.1) optimal trajectories with control $\tilde{u}$ lift to curves in the cotangent bundle satisfying $\dot{\lambda} = \vec{h}_{\tilde{u}(t)}(\lambda)$ and $\max_u h_u(\lambda) = h_{\tilde{u}}(\lambda)$.

Fix an optimal control $\tilde{u}$ and the relative extremal λ_t. Consider the function $h_{\tilde{u}}(\lambda) = h_{\tilde{u}(t)}(\lambda)$ and define the following non autonomous flow (which will play the role of parallel transport in this context):

$$\frac{d}{dt}\tilde{\Phi}_t = \vec{h}_{\tilde{u}}(\tilde{\Phi}_t) \qquad \tilde{\Phi}_0 = Id \tag{2.1}$$

It has the following properties:

i) It extends to the cotangent bundle the flow which solves $\dot{q} = f^t_{\tilde{u}}(q)$ on the base. In particular if λ_t is an extremal with initial condition λ_0, $\pi(\tilde{\Phi}_t(\lambda_0)) = q_{\tilde{u}}(t)$ where $q_{\tilde{u}}$ is an extremal trajectory.

ii) $\tilde{\Phi}_t$ preserves the fibre over each $q \in M$. The restriction $\tilde{\Phi}_t : T^*_qM \to T^*_{\tilde{\Phi}_t(q)}M$ is an affine transformation.

We use the symplectomorphism $\tilde{\Phi}_t$ to pull back the whole curve λ_t to the starting point λ_0. We can express all the first and second order information about the extremal using the following map and its derivatives:

$$b^t_u(\lambda) = (h^t_u - h^t_{\tilde{u}}) \circ \tilde{\Phi}_t(\lambda)$$

Notice that:

i) $b^t_u(\lambda_0)|_{u=\tilde{u}(t)} = 0 = d_{\lambda_0} b^t_u|_{u=\tilde{u}(t)}$ by definition.

ii) $\partial_u b^t_u|_{u=\tilde{u}(t)} = \partial_u(h^t_u \circ \tilde{\Phi}_t)|_{u=\tilde{u}(t)} = 0$ since $\lambda(t)$ is an extremal and $\tilde{u}$ the relative control.

Thus the first non zero derivatives are the order two ones. We define the following maps:

$$\begin{aligned} Z_t &= \partial_u \vec{b}^t_u(\lambda_0)|_{u=\tilde{u}(t)} : \mathbb{R}^k = T_{\tilde{u}(t)}U \to T_{\lambda_0}(T^*M) \\ H_t &= \partial^2_u b_t(\lambda_0)|_{u=\tilde{u}(t)} : \mathbb{R}^k = T_{\tilde{u}(t)}U \to T^*_{\tilde{u}(t)}U = \mathbb{R}^k \end{aligned} \tag{2.2}$$

We denote by $\Pi = \ker \pi_*$ the kernel of the differential of the natural projection $\pi : T^*M \to M$.

Proposition (Differential of the endpoint map)**.** *Consider the endpoint map $E^t : \mathcal{U}_{q_0} \to M$. Fix a point $\tilde{u}$ and consider the symplectomorphism $\tilde{\Phi}_t$ and the map Z_t defined above. The differential is the following map:*

$$d_{\tilde{u}}E(v_t) = d_{\lambda(t)}\pi \circ d_{\lambda_0}\tilde{\Phi}_t\left(\int_0^t Z_\tau v_\tau d\tau\right) \in T_{q_t}M$$

In particular, if we identify $T_{\lambda_0}(T^*M)$ with $\mathbb{R}^{2m}$ and write $Z_t = \begin{pmatrix} Y_t \\ X_t \end{pmatrix}$, $\tilde{u}$ is a regular point if and only if $v_t \mapsto \int_0^t X_\tau v_\tau d\tau$ is surjective. Equivalently if the following matrix is invertible:

$$\Gamma_t = \int_0^t X_\tau X_\tau^* d\tau \in Mat_{n\times n}(\mathbb{R}), \quad \det(\Gamma_t) \neq 0$$

If $d_{\tilde{u}}E^t$ is surjective then $(E^t)^{-1}(q_t)$ is smooth in a neighbourhood of $\tilde{u}$ and is tangent space is given by:

$$\begin{aligned} T_{\tilde{u}}(E^t)^{-1}(q_t) &= \{v \in L^\infty([0,1],\mathbb{R}^k) : \int_0^t X_\tau v_\tau d\tau = 0\} \\ &= \{v \in L^\infty([0,1],\mathbb{R}^k) : \int_0^t Z_\tau v_\tau d\tau \in \Pi\} \end{aligned}$$

In this case optimal control problems can be formulated as constrained minimization problems. We take as submanifold of $L^\infty([0,1],\mathbb{R}^k)$ the set $(E^t_{q_0})^{-1}(q_1)$ and functional to minimize the cost $\mathcal{J}$. In the sequel we will often restrict to this situation. See [14] for a more detailed discussion.

At critical points is well defined (i.e. independent of coordinates) the second variation of $d_{\tilde{u}}E^t_{q_0}$. It is a map defined on the kernel of the first differential with values in the cokernel (see appendix B or [12][Chapter 20] for further details). When the image of the first differential is an hyperplane it reduces to a quadratic form. Such extremals are called of *corank* 1, see Definition B.2. This happens, for instance, when we are in the situation of constrained minimization described above. In this case the second derivative of the extended Endpoint map (i.e. lifted to manifold $M \times \mathbb{R}$ to include the cost as a variable) can be expressed with the function Z_t and H_t defined in eq. (2.2).

Using chronological calculus (see again [12] or [5] and proposition B.2) it is possible to write the second derivative of the functional $\mathcal{J}$ (i.e. the second derivative of the extended Endpoint map) on $\ker d_{\tilde{u}}E^t \subseteq L^\infty([0,1],\mathbb{R}^k)$ as follows.

Proposition (Second variation). *Suppose that $(\lambda(t),\tilde{u})$ is a critical point of $\mathcal{J}$ with fixed initial and final point. For any $u \in \ker d_{\tilde{u}}E^t_{q_0}$ the second variation has the following expression:*

$$d^2_{\tilde{u}}\mathcal{J}(u) = -\int_0^1 \langle H_t u_t, u_t\rangle dt - \int_0^1\int_0^t \sigma(Z_\tau u_\tau, Z_t u_t) d\tau dt$$

The associated bilinear form is symmetric provided that u, v lie in a subspace that projects to a Lagrangian one via the map $u \mapsto \int_0^1 Z_t u_t dt$.

$$d^2_{\tilde{u}}\mathcal{J}(u) = -\int_0^1 \langle H_t u_t, v_t\rangle dt - \int_0^1\int_0^t \sigma(Z_\tau u_\tau, Z_t v_t) d\tau dt$$

The formula just derived for the second derivative of the cost in the case of an extremal which is also a *regular* point of the Endpoint map actually hold in more general context. In particular, for any $\eta \in T^*_{q_t}M$ such that $\eta d_{\tilde{u}}E^t_{q_0} = 0$ the second variation can always be written as:

$$\eta\, d^2_{\tilde{u}}E^t_{q_0}(u) = -\int_0^t \langle \tilde{H}_t u_t, u_t\rangle - \int_0^t\int_0^\tau \sigma_{\tilde{P}^*_t(\eta)}(\tilde{Z}_\tau u_\tau, \tilde{Z}_t u_t) d\tau dt$$

Where P_t denotes the flow of $f_{\tilde{u}(t)}(q)$ at time t, $\tilde{Z}_t u_t$ is an element of $T_{\tilde{P}_t^*\eta}T^*M$ and $\tilde{H}_t$ a family of symmetric matrices. In the following section we are going thus to consider abstractly operators on $L^\infty([0,1],\mathbb{R}^k)$ (or their extension to L^2) of this form, focusing especially on the compact part:

$$K(u)(t) = \int_0^1 \sigma(Z_t\cdot, Z_\tau u_\tau)d\tau. \tag{2.3}$$

2.2 Characterization of the Second Variation

This section is devoted to the proof of Theorem 1.3 and his reformulation in Theorem 1.4. It roughly says that any integral operators with integral kernel of the form given in eq. (1.6) which becomes symmetric on a finite codimension subspace actually can be represented as in eq. (2.3) and it is completely determined by its skew-symmetric part. We recall here the statement:

Theorem. *Let be K an operator satisfying eq.* (1.5) *and eq.* (1.6)*. Then $\mathcal{A}$ has finite rank and completely determines K. More precisely, if $\mathcal{A}$ has rank $2m$ and is represented as:*

$$\mathcal{A}(v)(t) := \frac{1}{2}Z_t^*\mathcal{A}_0 \int_0^1 Z_\tau v(\tau)dt,$$

for a skew-symmetric $2m \times 2m$ matrix $\mathcal{A}_0$ and a $2m \times k$ matrix Z_t then:

$$K(v)(t) = \int_0^t Z_t^*\mathcal{A}_0 Z_\tau v(\tau)d\tau. \tag{2.4}$$

Proof of Theorem 1.3. The proof of the first part of the statement follows from a couple of elementary considerations. In the sequel we will use the short-hand notation $\mathcal{A}$ for $Skew(K)$.

Fact 1: Equation (1.5) *holds if and only if $\mathcal{A}$ has finite rank*

Suppose that $K|_\mathcal{V}$ is symmetric. Consider the orthogonal splitting of $L^2[0,1]$ as $\mathcal{V}\oplus\mathcal{V}^\perp$. Equation (1.5) can be reformulated as $\mathcal{A}(\mathcal{V}) \subseteq \mathcal{V}^\perp$, thus $\operatorname{Im}(\mathcal{A}(L^2[0,1])) \subseteq \mathcal{V}^\perp + \mathcal{A}(\mathcal{V}^\perp)$ which is finite dimensional.

Conversely, if the range of $\mathcal{A}$ is finite dimensional, we can decompose $L^2[0,1]$ as $\operatorname{Im}(\mathcal{A})\oplus\ker(\mathcal{A})$, where the decomposition is orthogonal by skew-symmetry. Thus, on $\ker(\mathcal{A})$, K is symmetric.

Fact 2: $\mathcal{A}$ determines the kernel of K

It is well known that, if K is Hilbert-Schmidt, then K^* is Hilbert-Schmidt too. Since we are assuming eq. (1.6) it is given by:

$$K^*(v)(t) = \int_t^1 V^*(\tau,t)v(\tau)d\tau.$$

So we can write down the integral kernel $A(t,\tau)$ of $\mathcal{A}$ as follows:

$$A(t,\tau) = \begin{cases} \frac{1}{2}V(t,\tau) \text{ if } \tau < t \\ -\frac{1}{2}V^*(\tau,t) \text{ if } t < \tau. \end{cases}$$

The key observation now is that the support of the kernel of K is disjoint form the support of the kernel of K^*. Thus the kernel of $\mathcal{A}$ determines the kernel of K (and vice versa).

Now, since we are assuming that $\mathcal{A}$ has finite dimensional image, we can present its kernel as:

$$A(t,\tau) = \frac{1}{2} Z_t^* \mathcal{A}_0 Z_\tau,$$

where $\mathcal{A}_0$ is a skew-symmetric matrix and Z_t is a $\dim(\mathrm{Im}(\mathcal{A})) \times k$ matrix that has as rows the elements of some orthonormal base of $\mathrm{Im}(\mathcal{A})$. Without loss of generality we can assume $\mathcal{A}_0 = J$. In fact with an orthogonal change of coordinates $\mathcal{A}_0$ decomposes as a direct sum of rotation with an amplitude λ_i. Rescaling the coordinates by $\sqrt{\lambda_i}$ yields the desired canonical form J. □

Now we prove corollary 1.2. This result relates the spectrum of the skew-symmetric part of any operator satisfying eqs. (1.5) and (1.6) and its capacity, a spectral invariant introduced in [7] whose definition was recalled in Section 1.3.

Proof of Corollary 1.2. First of all notice that, now that we have written down any operator satisfying eqs. (1.5) and (1.6) in the same form as those in eq. (1.4), we can apply all the results about the asymptotic of their eigenvalues. In particular, if we assume that the space $\mathrm{Im}(\mathcal{A}) \subset L^2([0,1],\mathbb{R}^k)$ is generated by piecewise analytic functions, the ordered sequence of eigenvalues satisfies:

$$\lambda_n = \frac{\xi}{\pi n} + O(n^{-5/3}), \qquad \text{as } n \to \pm\infty.$$

Notice that we are using a better estimates on the reminder (for the case of the $1-$capacity) then the one given in Theorem 1.5 that was given in [7]. We denote by $M^\dagger = \bar{M}^*$ the conjugate transpose. Set $2m = \dim(\mathrm{Im}(\mathcal{A}))$, since the map $t \mapsto Z_t$ is analytic, there exists a piecewise analytic family of unitary matrices G_t such that:

$$G_t^\dagger Z_t^* J Z_t G_t = \begin{bmatrix} i\zeta_1(t) & & & & & & \\ & \ddots & & & & & \\ & & i\zeta_l(t) & & & & \\ & & & -i\zeta_1(t) & & & \\ & & & & \ddots & & \\ & & & & & -i\zeta_l(t) & \\ & & & & & & \underline{0} \end{bmatrix}$$

Without loss of generality we can assume that the function ζ_i are analytic on the whole interval and everywhere non negative. Recall that the coefficient ξ appearing in the asymptotic was computed as $\xi = \int_0^1 \zeta(t)dt = \int_0^1 \sum_{i=0}^l \zeta_i(t)dt$.

Let us work on the Hilbert space $L^2([0,1],\mathbb{C}^k)$ with standard hermitian product. Notice that $G : L^2([0,1],\mathbb{C}^k) \to L^2([0,1],\mathbb{C}^k)$, $v \mapsto G_t v$ is an isometry, thus the eigenvalue of $Skew(K) = \mathcal{A}$ remain the same if we consider the similar operator $G^{-1} \circ \mathcal{A} \circ G$ which acts as follows:

$$G^{-1} \circ \mathcal{A} \circ G(v) = \frac{1}{2} G_t^\dagger Z_t^* J \int_0^1 Z_\tau G_\tau v(\tau) d\tau$$

To simplify notation let's forget about this change of coordinates and still call Z_t the matrix Z_tG_t. Write Z_t as:

$$Z_t = \begin{pmatrix} y_1^*(t) \\ \vdots \\ y_m^*(t) \\ x_1^*(t) \\ \vdots \\ x_m^*(t) \end{pmatrix}.$$

We introduce the following notation: for a vector function v_i the quantity $(v_i)_j$ stands for $j-$th component of v_i.

We can now bound the function $\zeta(t)$ in terms of the components of the matrix Z_t:

$$\begin{aligned} 2\zeta(t) &= \sum_{j=1}^{k} |(Z_t^\dagger J Z_t)_{jj}| \leq \sum_{i=1}^{m}\sum_{j=1}^{k} |(x_i)_j(\bar{y}_i)_j - (y_i)_j(\bar{x}_i)_j|(t) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{k} 2|\mathrm{Im}((x_i)_j(\bar{y}_i)_j)| \leq \sum_{i=1}^{m}\sum_{j=1}^{k} 2|(x_i)_j||(y_i)_j| = \sum_{i=1}^{m} 2\langle |x_i|, |y_i|\rangle(t) \end{aligned}$$

Where the vector $|v|$ is the vector with entries the absolute values the entries of v. Integrating and using Hölder inequality for the 2 norm, we get:

$$\xi = \int_0^1 \zeta(t)dt = \sum_{i=1}^{m} ||x_i||_2\, ||y_i||_2.$$

The next step is to relate the quantity on the right hand side to the eigenvalues of $\mathcal{A}$. The strategy now is to modify the matrix Z_t in order to get an orthonormal frame of $Im(\mathcal{A})$. Keeping track of the transformations used we get a matrix representing $\mathcal{A}$, then it is enough to compute the eigenvalues of the said matrix.

We can assume, without loss of generality that $\langle x_i, x_j\rangle_{L^2} = \delta_{ij}$. This can be achieved with a symplectic change of the matrix Z_t. Then we modify the y_j in order to make them orthogonal to the space generated by the x_j. We use the following transformation:

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} \mapsto \begin{pmatrix} 1 & M \\ 0 & 1 \end{pmatrix}\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} Y_t + MX_t \\ X_t \end{pmatrix}$$

where M is defined by the relation $\int_0^1 Y_tX_t^* + MX_tX_t^* dt = \int_0^1 Y_tX_t^*dt + M = 0$. The last step is to make y_j orthonormal. If we multiply Y_t by a matrix L we find the equation $L\int_0^1 Y_tY_t^*dtL^* = 1$, so $L = (\int_0^1 Y_tY_t^*dt)^{-\frac{1}{2}}$. Thus the matrix representing $\mathcal{A}$ in this coordinates is one half of:

$$\mathcal{A}_0 = \begin{pmatrix} L^{-1} & 0 \\ -M^* & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} L^{-1} & -M \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & L^{-1} \\ -L^{-1} & M^* - M \end{pmatrix}$$

If we square $\mathcal{A}_0$ and compute the trace we get:

$$-\frac{1}{2}\,\mathrm{tr}(\mathcal{A}_0^2) = \mathrm{tr}(L^{-2}) - \frac{1}{2}\,\mathrm{tr}((M^* - M)^2) \geq \mathrm{tr}\left(\int_0^1 Y_tY_t^*dt\right) = \sum_{i=1}^{m} ||y_i||_2^2$$

Call $\Sigma(\mathcal{A})$ the spectrum of $\mathcal{A}$, since $\mathcal{A}$ is skew-symmetric it follows that:

$$-\frac{1}{2}\operatorname{tr}(\mathcal{A}_0^2) = 4 \sum_{\mu\in\Sigma(\mathcal{A}),-i\mu>0} -\mu^2 \geq 0.$$

Recalling that $||x_i|| = 1$ and putting all together we find that:

$$\xi \leq \sum_{i=1}^m ||y_i||_2 \leq \sqrt{m}\sqrt{\sum_{i=1}^m ||y_i||_2^2} = 2\sqrt{m}\sqrt{\sum_{\mu\in\Sigma(\mathcal{A}),-i\mu>0} -\mu^2}.$$

□

Example 2.1. Consider a matrix Z_t of the following form:

$$Z_t = \begin{bmatrix} \xi_1(t) & \xi_3(t) \\ 0 & \xi_2(t) \end{bmatrix} \quad Z_t^* J Z_t = \begin{bmatrix} 0 & -\xi_1\xi_2(t) \\ \xi_2\xi_1(t) & 0 \end{bmatrix}$$

The capacity of K is given by $\zeta = \int_0^1 |\xi_1\xi_2|(t)dt$. We can assume that $\langle \xi_2, \xi_3\rangle = 0$ and $||\xi_2|| = 1$.

A direct computation shows that the eigenvalue of $SkewK$ are $\frac{\pm i}{2}\sqrt{(||\xi_1||^2 + ||\xi_3||^2)}$. This shows that the two quantities behave in a very different way. If we choose ξ_2 very close to ξ_1 and ξ_3 small, capacity and eigenvalue square are comparable. If we choose ξ_3 very big the capacity remains the same whereas the eigenvalues explode. In particular there cannot be any lower bound of ζ in terms of the eigenvalues of K.

Remark 2.1. There is a natural class of translations that preserves the capacity. Take any path Φ_t of symplectic matrices (say L^2 integrable), the operators constructed with Z_t and $\Phi_t Z_t$ have the same capacity (but the respective skew-symmetric part clearly do not have the same eigenvalues).

Set $K^\Phi(v) = \int_0^t Z_t^* J \Phi_t^{-1}\Phi_\tau Z_\tau v_\tau d\tau$ and $\Sigma^+(K^\Phi)$ the set of eigenvalues of $Skew(K^\Phi)$ satisfying $-i\sigma \geq 0$. It seems natural to ask if:

$$\zeta(K) = 2 \inf_{\Phi_t\in Sp(n)} \sqrt{\sum_{\sigma\in\Sigma^+(K^\Phi)} -\sigma^2}$$

Take for instance the example above and suppose for simplicity that ξ_1 and ξ_2 are positive and never vanishing. Using the following transformation we obtain:

$$Z_t' = \begin{bmatrix} \sqrt{\frac{\xi_2}{\xi_1}} & \frac{-\xi_3}{\sqrt{\xi_1\xi_2}} \\ 0 & \sqrt{\frac{\xi_1}{\xi_2}} \end{bmatrix} \begin{bmatrix} \xi_1 & \xi_3 \\ 0 & \xi_2 \end{bmatrix} = \begin{bmatrix} \sqrt{\xi_1\xi_2} & 0 \\ 0 & \sqrt{\xi_1\xi_2} \end{bmatrix}$$

and in this case the eigenvalue became $\frac{\pm i}{2}\langle \xi_1, \xi_2\rangle$, precisely half the capacity.

2.2.1 Proof Theorem 1.4

In this section we reformulate Theorem 1.3 as a characterization of the compact part of the second variation of an optimal control problem at a strictly normal regular extremal (see definition B.2). The statement we are going to prove is the following:

Theorem. *Suppose $\mathcal{V} \subset L^2([0,1],\mathbb{R}^k)$ is a finite codimension subspace and K and operator satisfying eqs. (1.5) and (1.6). Then $(K,\mathcal{V})$ can be realized as the second variation of an optimal control problem at a strictly normal regular extremal. To any such couple we can associate a triple $((\Sigma,\sigma),\Pi,Z)$ consisting of:*

i) a finite dimensional symplectic space (Σ,σ);

ii) a Lagrangian subspace $\Pi \subset \Sigma$;

iii) a linear map $Z : L^2([0,1],\mathbb{R}^k) \to \Sigma$ such that $\mathrm{Im}(Z)$ is transversal to the subspace Π.

This triple is unique up to the action of $\mathrm{stab}_\Pi(\Sigma,\sigma)$, the group of symplectic transformations that fix Π. Any other triple is given by $((\Sigma,\sigma),\Pi,\Phi\circ Z)$ for $\Phi \in \mathrm{stab}_\Pi(\Sigma,\sigma)$.

Vice versa any triple $((\Sigma,\sigma),\Pi,Z)$ as above determines a couple $(K,\mathcal{V})$. We can define the skew-symmetric part $\mathcal{A}$ of K as:

$$\langle \mathcal{A}u, v\rangle = \sigma(Zu, Zv),\ \forall u,v \in L^2([0,1],\mathbb{R}^k),$$

$\mathcal{A}$ determines the whole operator K and its domain is recovered as $\mathcal{V} = Z^{-1}(\Pi)$.

Remark 2.2. The uniqueness part of the theorem maybe needs a quick comment. One should think of this indeterminacy as the dependence of the second variation on the choice of a particular set of coordinates on the base manifold. Choosing coordinates at a point λ_0 is equivalent to choosing a Lagrange subspace complementary to Π. The action of $\mathrm{stab}_\Pi(\Sigma,\sigma)$ is transitive on $\Pi^\pitchfork$.

Proof. The proof is essentially a reformulation of Theorem 1.3. Given the operator we construct the symplectic space (Σ,σ) taking as vector space the image of the skew-symmetric part $\mathrm{Im}(\mathcal{A})$ and as symplectic form $\langle \mathcal{A}\cdot,\cdot\rangle$.

The transversality condition correspond to the fact that the differential of the endpoint map is surjective.

The only thing left to show is uniqueness of the triple. Without loss of generality we can assume that the symplectic subspace $(\Sigma,\sigma) = (\mathbb{R}^{2n},\sigma)$ is the standard one and that the Lagrangian subspace Π is the vertical subspace. In this coordinates

$$Z(v) = \int_0^1 Z_t v_t dt = \int_0^1 \begin{pmatrix} Y_t \\ X_t \end{pmatrix} v_t dt.$$

Define the following map:

$$F : L^2([0,1],\mathrm{Mat}_{n\times k}(\mathbb{R})) \to L^2([0,1]^2,\mathrm{Mat}_{k\times k}(\mathbb{R})),\quad Y_t \mapsto Z_t^* J Z_\tau = X_t^* Y_\tau - Y_t^* X_\tau.$$

It is linear if X_t is fixed. To determine uniqueness we have to study an affine equation thus is sufficient to study the kernel of F. Suppose for simplicity that X_t and Y_t are continuous in t. We have to solve the equation:

$$F(Y_t) = Z_t^* J Z_\tau = \sigma(Z_t, Z_\tau) = 0.$$

Consider the following subspace of $\mathbb{R}^{2n}$

$$V^{[0,1]} = \left\{ \sum_{i=1}^{l} Z_{t_i}\nu_i : \nu_i \in \mathbb{R}^k, t_i \in [0,1], l \in \mathbb{N} \right\} \subset \mathbb{R}^{2n}$$

It follows that $F(Y_t) = 0$ if and only if the subspace $V^{[0,1]}$ is isotropic. Since we are in finite dimension, we can consider a finite number of instants t_i to which we can restrict to generate the whole $V^{[0,1]}$. Call I the set of this instants. Without loss of generality we can assume that $\{\sum_{i\in I} X_{t_i}\nu_i, \nu_i \in \mathbb{R}^k, t_i \in I\} = \mathbb{R}^n$.

This is so since the image of Z is transversal to Π and thus $\Gamma = \int_0^1 X_t X_t^* dt$ is non degenerate. In fact, if the subspace $\{\sum_{i=1}^l X_{t_i}\nu_i | \, \nu_i \in \mathbb{R}^k, l \in \mathbb{N}\}$ were a proper subspace of $\mathbb{R}^n$, there would be a vector μ such that $\langle \mu, X_t\nu\rangle = 0$, $\forall t \in [0,1]$ and $\forall \nu \in \mathbb{R}^n$. Thus an element of the kernel of Γ. A contradiction.

Now we evaluate the equation $F(Y_t) = 0 \iff Y_t^* X_\tau = X_t^* Y_\tau$ at the instants $t = t_i$ that guarantee controllability. One can read off the following identities:

$$Y_t^* v_j = X_t^* c_j$$

where the v_j's are a base of $\mathbb{R}^n$ and c_j free parameters. Taking transpose we get that $Y_t = GX_t$.

It is straightforward to check that, if $Y_t = GX_t$, G must be symmetric, in fact:

$$Z_t J Z_\tau = Y_t^* X_\tau - X_t^* Y_\tau = X_t^*(G^* - G)X_\tau = 0 \iff G = G^*$$

And so uniqueness is proved when X_t and Y_t are continuous.

The case in which X_t and Y_t are just L^2 (matrix-)functions can be dealt with similarly. One has just to replace *evaluations* with integrals of the form $\int_{t-\epsilon}^{t+\epsilon} Z_\tau \nu d\tau$ and $\int_{t-\epsilon}^{t+\epsilon} X_\tau \nu d\tau$ and interpret every equality t almost everywhere.

The only thing left to show is how to construct a control system with given $(K, \mathcal{V})$ as second variation. By the equivalence stated above it is enough to show that we can realize any given map $Z : L^2([0,1], \mathbb{R}^k) \to \Sigma$ with a proper control system. We can assume without loss of generality that (Σ, σ) is just $\mathbb{R}^{2m}$ with the standard symplectic form and Π is the vertical subspace. With this choices the map Z is given by :

$$v \mapsto \int_0^1 Z_t v_t dt = \int_0^1 \begin{pmatrix} Y_t v_t \\ X_t v_t \end{pmatrix} dt$$

The operator K is then given by $K(v) = \int_0^t Z_t^* J Z_\tau v_\tau d\tau$ and $\mathcal{V} = \{v | \int_0^1 X_t v_t dt = 0\}$. Consider the following linear quadratic system on $\mathbb{R}^m$:

$$f_u(q) = B_t u \quad \varphi_t(x) = \frac{1}{2}|u|^2 + \langle \Omega_t u, x\rangle,$$

where B_t and Ω_t are matrices of size $m \times k$, the Hamiltonian in PMP reads:

$$h_u(\lambda, x) = \langle \lambda, B_t u\rangle - \frac{1}{2}|u|^2 - \langle \Omega_t u, x\rangle$$

Take as extremal control $u_t \equiv 0$, it easy to check that the re-parametrization flow $\tilde{\Phi}_t$ defined in eq. (B.11) is just the identity and the matrix Z_t for this problem is the following:

$$Z_t = \begin{pmatrix} \Omega_t \\ B_t \end{pmatrix}$$

So it is enough to take $\Omega_t = Y_t$ and $B_t = X_t$. □

2.3 Second Variation with moving endpoints

Up until now we have consider just optimal control problems with Dirichlet boundary condition. In this section we construct a suitable space of variations for optimal control problems with boundary conditions $N \subseteq M \times M$ which will be always used in the sequel. The idea behind the construction is rather simple: we wish to employ the machinery introduce in the previous section to differentiate the Endpoint map in the case of Dirichlet boundary condition. To do so we introduce now a auxiliary control system that extends our original one and incorporates the boundary conditions. We will consider at first the case of separated boundary conditions, i.e. N of the form $N = N_0 \times N_1$.

Assume that we are given a control system as in eq. (B.2) and an extremal λ_t satisfying the first order optimality condition of PMP. Denote by λ_0 and λ_1 the initial and final covector of said extremal and by q_0 and q_1 the respective projection on the base manifold.

Let us fix a family of vector fields $X_0^1, \dots, X_0^{\dim(N_0)}$ and $X_1^j, \dots X_1^{\dim(N_1)}$ with the following properties:

i) $\operatorname{span}\{X_i^j\}_{j=1}^{\dim(N_i)} = T_q N_i$ for any q in a neighbourhood (inside N_i) of q_i.

ii) they generate a local integrable distribution in a neighbourhood (inside M) of q_i. In particular the leaf through q_i coincide with N_i.

Consider now the following (linear with respect to the control) vector fields:

$$f_v^i(q) = \sum_{j=1}^{\dim(N_i)} X_i^j(q) v_j, \quad v \in \mathbb{R}^{\dim(N_i)}.$$

We take as controls $\hat{u}$ elements of the Hilbert space $\mathbb{R}^{\dim(N_0)} \oplus L^2([0,1], \mathbb{R}^k) \oplus \mathbb{R}^{\dim(N_1)}$. With a slight abuse of notation, we can think to this vectors as *functions* on $[-1,2]$ which are constant on $[0,1]^c$. We will denote by u_0 the value of $\hat{u}$ on $[-1,0)$, by u_1 its value on $(1,2]$ and by u the restriction on $[0,1]$. Define then the following optimal control problem:

$$\hat{f}_{\hat{u}}^t(q) = \begin{cases} f_{u_0}^0(q) \text{ for } t \in [-1,0), \\ f_u^t(q) \text{ for } t \in [0,1], \\ f_{u_1}^1(t) \text{ for } t \in (1,2]. \end{cases} \qquad \mathcal{J}(\hat{u}) = \int_0^1 \varphi(t, u(t), q_u(t)) dt \tag{2.5}$$

We have the following easy observation:

Lemma 2.1. *The problem given in eq.* (B.2) *with moving boundary conditions* $N = N_0 \times N_1$ *is locally equivalent to the one in eq.* (2.5) *with fixed endpoints.*

Proof. Recall that we are fixing an extremal λ_t and call $\tilde{u}$ the relative optimal control. It is also an extremal for the auxiliary system, with optimal control given by $\bar{u}$. We can identify it with:

$$\hat{\lambda}_t = \begin{cases} \lambda_0 \text{ for } t \in [-1,0], \\ \lambda_t \text{ for } t \in [0,1], \\ \lambda_1 \text{ for } t \in [1,2]. \end{cases} \qquad \hat{u} = \begin{cases} 0 \text{ for } t \in [-1,0), \\ \tilde{u} \text{ for } t \in [0,1], \\ 0 \text{ for } t \in (1,2]. \end{cases} \tag{2.6}$$

For any variation of the curve $q_t = \pi\lambda_t$ with endpoints, say (r_0, r_1), sufficiently close to (q_0, q_1) there exists controls (u_0, u_1) such that:

$$\begin{cases} \dot{q} = f_{u_0}(q), \\ q(-1) = q_0, \\ q(0) = r_0. \end{cases} \qquad \begin{cases} \dot{q} = f_{u_1}(q), \\ q(1) = r_1, \\ q(2) = q_1. \end{cases}$$

It follows that q_t is also a variation for the auxiliary system. Vice-versa, since we fixed an integrable distribution and assumed that the leaf through (q_0, q_1) coincides with $N_0 \times N_1$, any integral curve of the fields f_v^0 and f_v^1 does not leave N_0 nor N_1 respectively. Thus, the restriction of the variation to $[0, 1]$ gives a variation for the original system with the right boundary conditions. The situation is depicted in fig. 2.1.

The last observation we have to make to conclude is that the cost functional is the same for both the problems. □

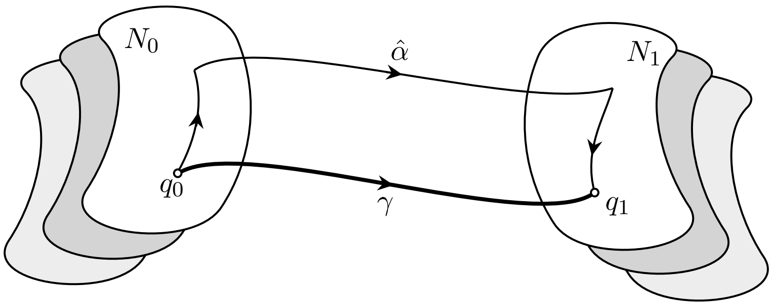


Figure 2.1: An admissible extended variation $\hat{\alpha}$ of an extremal curve γ

The curve described in eq. (2.6) is actually an extremal for the auxiliary problem. We are going to write down the first order conditions coming from PMP and check this. The Hamiltonians appearing in PMP read:

$$h^t_{\hat{u}}(\lambda) = \begin{cases} \langle \lambda, f^0_{u_0}(q) \rangle \text{ for } t \in [-1, 0), \\ \langle \lambda, f^t_u(q) \rangle - \varphi(t, u, q) \text{ for } t \in [0, 1], \\ \langle \lambda, f^1_{u_1}(q) \rangle \text{ for } t \in (1, 2]. \end{cases}$$

In particular the maximum condition $\max_v h^t_v(\hat{\lambda}_t) = h^t_{\hat{u}}(\hat{\lambda}_t)$ implies that the initial condition (λ_0, λ_1) must lie in the annihilator of the boundary conditions. This is because a linear function has maximum only if it is constantly zero.

Now we can differentiate the Endpoint map of the auxiliary system using the machinery for fixed endpoints and get integral expression for the first and second variation. Let us start with the flow generated by $\vec{h}^t_{\hat{u}}$. By the discussion above we have that (λ_0, λ_1) lie in the annihilator of $N_0 \times N_1$ and thus they are equilibrium points of the $\vec{h}^t_{\hat{u}}$. The flow reads:

$$\hat{\Phi}_t = \begin{cases} 1 & \text{for } t \in [-1, 0), \\ \tilde{\Phi}_t & \text{for } t \in [0, 1], \\ \tilde{\Phi}_1 & \text{for } t \in (1, 2]. \end{cases}$$

So we can simplify the expression of $\vec{b}^t_{\hat{u}}(\lambda) = (\vec{h}^t_{\hat{u}} - \vec{h}^t_{\bar{u}}) \circ \hat{\Phi}_t(\lambda)$ and of its derivatives at $\lambda = \lambda_0$. We have the following identities:

$$\vec{b}^t_{\hat{u}}(\lambda) = \begin{cases} \overrightarrow{\langle \lambda, f^0_{u_0}(q) \rangle} & \text{for } t \in [-1, 0), \\ \vec{b}^t_u(\lambda) & \text{for } t \in [0, 1], \\ \overrightarrow{\langle \tilde{\Phi}^*_1 \lambda, f^1_{u_1}(q) \rangle} & \text{for } t \in (1, 2]. \end{cases}$$

$$\hat{Z}_t = \partial_{\hat{u}} \vec{b}^t_{\hat{u}}(\lambda)|_{\lambda=\lambda_0} = \begin{cases} Z_0 & \text{for } t \in [-1, 0), \\ Z_t & \text{for } t \in [0, 1], \\ Z_1 & \text{for } t \in (1, 2]. \end{cases} \qquad \hat{H}_t = \begin{cases} 0 & \text{for } t \in [-1, 0), \\ H_t & \text{for } t \in [0, 1], \\ 0 & \text{for } t \in (1, 2]. \end{cases}$$

In this case the maps Z_0 and Z_1 are constant, the next lemma shows that they take value in the annihilators of the boundary conditions:

Lemma 2.2. *The image of $Z_0 : \mathbb{R}^{\dim(N_0)} \to T_{\lambda_0} T^* M$ is contained in $T_{\lambda_0} Ann(N_0)$. Similarly $(\tilde{\Phi}_1)_* Z_1$ maps $\mathbb{R}^{\dim(N_1)}$ into the tangent space of $Ann(N_1)$. Both maps are injective and transversal to $\Pi_{\lambda_i} \cap T_{\lambda_i} Ann(N_i)$.*

Proof. Recall that f_{0i} generate the tangent space to N_0 close to q_0. Define the Hamiltonians

$$l_j(\lambda) = \langle \lambda, X^0_j \rangle, \qquad j = 1, \dots, \dim N_0.$$

Then $Ann(N_0)$ can be equivalently described as the common part of the zero locus of l_j inside $\pi^{-1}(N_0)$:

$$Ann(N_0) = \{\lambda \in T^*M : \pi(\lambda) \in N_0,\, l_j(\lambda) = 0,\, j = 1, \dots, \dim N_0\}.$$

But then by the definition of a Hamiltonian vector field

$$d_{\lambda_0} l_j(Z_0 v) = d_{\lambda_0} l_j(\vec{l}_k v_k) = \sigma_{\lambda_0}(\vec{l}_k v_k, \vec{l}_i) = \langle \lambda_0, [X^0_k v_k, X^0_i] \rangle = 0,$$

where the last equality is due to involutivity of the family $X^0_k v_k$.

Similarly one has that $Z_1 u_1$ is always tangent to the image of $Ann(N_1)$ under the differential of $(\tilde{\Phi}_1)^{-1}$. □

Following the discussion of Appendix B, we have that the differential of the Endpoint map at the point $\hat{u}$ is given by:

$$d_{\hat{u}} E(\hat{v}) = \pi_* \int_{-1}^{2} \hat{Z}_t \hat{v}_t dt = \pi_* \left(Z_0 v_0 + \int_0^1 Z_t v_t dt + Z_1 v_1 \right).$$

In particular the tangent space to fixed endpoint variation is given by $\ker d_{\bar{u}} E$. We will call this subspace $\mathcal{V}$ and it is given by the following condition:

$$\mathcal{V} = \{(u_0, u, u_1) : Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1 \in \Pi_{\lambda_0}\}$$

We are now in the position to give the following definition:

Definition 2.1 (Second Variation)**.** The second variation at $\hat{u}$ is the quadratic form defined on $\mathcal{V}$:

$$\begin{aligned} Q(\hat{v}) = \int_0^1 \left(-\langle H_t v_t, v_t\rangle + \sigma(Z_t v_t, \int_0^t Z_\tau v_\tau d\tau + Z_0 v_0) \right) dt + \\ + \sigma(Z_0 v_0 + \int_0^1 Z_t v_t dt, Z_1 v_1). \end{aligned} \tag{2.7}$$

The formula in eq. (2.7) is obtain from eq. (B.9) and proposition B.2 using the fact that $\hat{Z}_t$ is locally constant with values in isotropic subspaces for $t \in [-1, 0] \cup [1, 2]$.

Up until now, we have consider boundary conditions of the type $N = N_0 \times N_1$. To the general case $N \subseteq M \times M$ we duplicate the variables and work on $M \times M$ with boundary conditions $\tilde{N} = \Delta \times N$. Here $\Delta \subseteq M \times M$ denotes the diagonal subspace given by $\Delta = \{(q, q) : q \in M\}$. The new dynamic is given by:

$$\tilde{f}_u^t(r, q) = \begin{pmatrix} 0 \\ f_u^t(q) \end{pmatrix}$$

It is straightforward to check that a curve satisfies the boundary conditions for the initial problem if and only if it satisfies the boundary conditions $\Delta \times N$ for the new one.

Remark 2.3. One could be more rigorous and work with actual functions living in a subspace of $L^2([-1, 2], \mathbb{R}^{k'})$ where $k' = \max\{\dim(N_0), \dim(N_1), k\}$ just redefining in the obvious way the various objects appearing in the formulas. Alternatively one can work with k non commuting vector fields at the endpoints and select a $\dim(N_i)$ subspace of controls that maps onto the tangent space of N_i. However, this just introduces notational nuisances and no tangible benefits.

Chapter 3

Asymptotics of the Spectrum

The focus of this chapter is the spectrum of the Second Variation. We were able to depict a quite complete picture and compute the asymptotic distribution of the eigenvalues of the Second Variation in a good degree of generality. The result presented in this chapter is contained in [18]. The starting point of this work was Theorem 1 of [7]. It gives an asymptotic relation for the order sequence of the eigenvalues of the compact part K of the second variation given in eq. (2.3). Namely if $\{\lambda_n\}_{n\in\mathbb{Z}}$ is said sequence we have either:

$$\lambda_n \sim \frac{\xi}{n}, n \to \pm\infty \text{ as } \text{ for some } \xi > 0 \quad \text{ or } \quad \lambda_n = O(n^{-2}).$$

Here ξ, called *capacity* or $1-$capacity in Definition 1.2, is determined explicitly by the integral kernel of K as explained in Section 1.3. The natural question which arose was: what happens when $\lambda_n = O(n^2)$? Can we provide a similar type of asymptotics? The answer turn out to be positive and it is given by Theorem 1.5.

The main difficulty we had to address is that we could not reduce the boundary value problem for the eigenvalues of K to a known one (or at least known to the author) like it was done in [7]. For this reason the proof of Theorem 1.5 is rather long and requires analysing explicitly various toy models to build proper estimates.

For reader's convenience we recall here the statement of Theorem 1.5; we will consider the following quadratic form on $L^2([0,1],\mathbb{R}^k)$:

$$\mathcal{Q}(v) = \langle v, K(v)\rangle = \int_0^1 \int_0^t \langle Z_t v(t), J Z_\tau v(\tau)\rangle d\tau dt.$$

We will always make the assumption that Z_t is a $2n\times k$ matrix which depends piecewise analytically on the parameter $t \in [0,1]$. For $j \geq 1$ recall the definition of the matrices $A_j(t)$, they are given as follows:

$$A_j(t) = \begin{cases} \big(Z_t^{(k)}\big)^* J Z_t^{(k)} & \text{if } j = 2k-1 \\ \big(Z_t^{(k-1)}\big)^* J Z_t^{(k)} & \text{if } j = 2k \end{cases}$$

We use ρ_t to denote any eigenvalue of the matrix $A_j(t)$. If $j = 2k$, define:

$$\mu^+_{t,2k} := \sum_{\rho_t : \rho_t > 0} \sqrt[2k]{\rho_t} \qquad \mu^-_{t,2k} := \sum_{\rho_t : \rho_t < 0} \sqrt[2k]{|\rho_t|}.$$

For odd indices, A_{2k-1} is skew-symmetric and thus the spectrum is purely imaginary. So we define the function:

$$\mu_{t,2k-1} = \sum_{\rho_t : -i\rho_t > 0} \sqrt[2k-1]{-i\rho_t}.$$

Theorem. *Let Q be the quadratic form defined in eq. (1.7). It has either $\infty-$capacity or $j-$capacity with remainder of order $\nu = 1/2$. More precisely, let $j \geq 1$ be the lowest integer such that $A_j(t)$ is not identically zero, then*

i) if $j = 2k-1$, the $(2k-1)-$capacity ξ is given by:

$$\xi = \left(\int_0^1 \mu_{t,2k-1} dt \right)^{2k-1},$$

and thus for $n \in \mathbb{Z}$ sufficiently large:

$$\lambda_n = \frac{\left(\int_0^1 \mu_{t,2k-1} dt \right)^{2k-1}}{(\pi n)^{2k-1}} + O(n^{-2k+1/2}).$$

ii) if $j = 2k$, the $2k-$capacity (ξ_+, ξ_-) is given by:

$$\xi_{\pm} = \left(\int_0^1 \mu^{\pm}_{t,2k} dt \right)^{2k},$$

and thus for $n \in \mathbb{Z}$ sufficiently large:

$$\lambda_n = \frac{\left(\int_0^1 \mu^{\pm}_{t,2k} dt \right)^{2k}}{(\pi n)^{2k}} + O(n^{-2k-1/2}).$$

iii) if $A_j(t) \equiv 0$ for any j then Q has $\infty-$capacity.

Before diving into the proof let me just remark that the asymptotic relation we prove in the statement apply to the Second Variation for optimal control problems with moving endpoints given in Definition 2.1 too. It is easy to check that after restricting it to a finite codimension subspace it coincides with the fixed points one and by Proposition 3.3 this does not alter the asymptotic relations of Theorem 1.5.

3.1 The principal term of Q

We begin now the proof Theorem 1.5. We start with Lemma 3.1 to single out the main contributions to the asymptotic of the eigenvalues of Q (the quadratic form defined in eq. (1.7)). The first non zero term of the decomposition we give will determine the rate of decaying of the eigenvalues (see Proposition 3.4).

Before showing this and prove the precise estimates we need to carry out the explicit computation of the asymptotic in some model cases, namely when the matrices A_j are constant. Then we have to show how the $j-$capacity behaves with respect to natural operations such as direct sum of quadratic form or restriction to finite codimension subspaces (Proposition 3.3).

Let us start with some notation:

$$v_k(t) = \int_0^t v_{k-1}(\tau)d\tau, \quad v_0(t) = v(t) \in L^2([0,1],\mathbb{R}^m)$$

Suppose that the map $t \mapsto Z_t$ is real analytic (or at least regular enough to perform the necessary derivatives) and integrate by parts twice:

$$\begin{aligned}
\mathcal{Q}(v) &= \int_0^1 \langle Z_t v(t), \int_0^t JZ_\tau v(\tau)d\tau\rangle dt \\
&= \int_0^1 \langle Z_t v(t), JZ_t v_1(t)\rangle - \langle Z_t v(t), \int_0^t J\dot{Z}_\tau v_1(\tau)d\tau\rangle dt \\
&= \int_0^1 \langle Z_t v(t), JZ_t v_1(t)\rangle + \langle Z_t v_1(t), J\dot{Z}_t v_1(t)\rangle dt + \\
&\quad + \int_0^1 \langle \dot{Z}_t v_1(t), J\int_0^t \dot{Z}_\tau v_1(\tau)d\tau\rangle dt - \Big[\langle \int_0^1 Z_t v(t)dt, J\int_0^1 \dot{Z}_t v_1(t)dt\rangle\Big]
\end{aligned}$$

If we impose the condition $\int_0^1 v_t dt = 0$ ($\iff$ $v_1(1) = 0$) the term in brackets vanishes:

$$\langle \int_0^1 Z_t v(t)dt, J\int_0^1 \dot{Z}_t v_1(t)dt\rangle = \langle \int_0^1 Z_t v(t)dt, JZ_1 v_1(1)\rangle - \langle \int_0^1 Z_t v(t)dt, J\int_0^1 Z_t v(t)dt\rangle$$

and we can write $\mathcal{Q}$ as a sum of three terms

$$\mathcal{Q}(v) = \mathcal{Q}_1(v) + \mathcal{Q}_2(v) + R_1(v)$$

In analogy we can make the following definitions:

$$\begin{aligned}
\mathcal{Q}_{2k-1}(v) &= \int_0^1 \langle Z_t^{(k-1)} v_{k-1}(t), JZ_t^{(k-1)} v_k(t)\rangle = \int_0^1 \langle v_{k-1}(t), A_{2k-1}(t)v_k(t)\rangle \\
\mathcal{Q}_{2k}(v) &= \int_0^1 \langle Z_t^{(k-1)} v_k(t), JZ_t^{(k)} v_k(t)\rangle dt = \int_0^1 \langle v_k(t), A_{2k}(t)v_k(t)\rangle dt \\
R_k &= \int_0^1 \langle Z_t^{(k)} v_k(t), J\int_0^t Z_\tau^{(k)} v_k(\tau)d\tau\rangle dt \\
V_k &= \{v \in L^2([0,1],\mathbb{R}^m) : v_l(1) = 0, \, \forall\, 0 < l \leq k\}
\end{aligned}$$

Here the matrices $A_j(t)$ are exactly those defined in eq. (1.8).

Lemma 3.1. *For every $j \in \mathbb{N}$, on the subspace V_j, the form $\mathcal{Q}$ can be represented as*

$$\mathcal{Q}(v) = \sum_{k=1}^{2j} \mathcal{Q}_k(v) + R_j(v) \tag{3.1}$$

The matrices $A_{2k}(t)$ are symmetric provided that $\frac{d}{dt}A_{2k-1}(t) \equiv 0$. On the other hand A_{2k-1} is always skew symmetric.

Proof. It is sufficient to notice that $R_1(v)$ has the same form as $\mathcal{Q}(v)$ but with v_1 instead of v and $\dot{Z}_t$ instead of Z_t. Thus the same scheme of integration by parts gives the decomposition.

Notice that $A_{2k}(t) = A_{2k}^*(t) + \frac{d}{dt}A_{2k-1}(t)$ thus the skew-symmetric part of $A_{2k}(t)$ is zero if A_{2k-1} is zero or constant. $A_{2k-1}(t)$ is always skew-symmetric by definition. $\square$

3.2 The asymptotic in some model cases

Now we would like to compute explicitly the spectrum of the $\mathcal{Q}_j$ when the matrices A_j are constant. Unfortunately describing the spectrum with boundary conditions given by the V_j is quite hard. Already for $\mathcal{Q}_4$ the equation determining it cannot be solved explicitly.

We will derive the Euler-Lagrange equation for $\mathcal{Q}_j$ and turn instead to periodic boundary conditions for which everything becomes very explicit and show how to relate the solution for the two boundary value problems we are considering. Let us write down the Euler-Lagrange equations for the forms $\mathcal{Q}_j$. If $j = 2k$ integration by parts yields:

$$\begin{aligned}\mathcal{Q}_{2k}(v) - \lambda||v||^2 &= \int_0^1 \langle v_k(t), A_{2k}v_k(t)\rangle - \lambda\langle v_0(t), v_0(t)\rangle dt \\ &= \int_0^1 \langle v_0(t), (-1)^k A_{2k}v_{2k}(t) - \lambda v_0(t)\rangle dt + \\ &\quad + \sum_{r=0}^{k-1}(-1)^r \Big[\langle v_{k-r}(t), A_{2k}v_{k+r+1}(t)\rangle\Big]_0^1\end{aligned}$$

Notice that the boundary terms vanish identically if we impose the vanishing of v_j for $1 \leq j \leq k$ at boundary points.

We change notation and define $w(t) = v_{2k}(t)$ and $w^{(j)}(t) = \frac{d^j}{dt^j}(w(t))$. The new equations are:

$$w^{(2k)}(t) = \frac{(-1)^k}{\lambda}A_{2k}w(t)$$

We can perform a linear change of coordinates that diagonalizes A_{2k} to reduce to m systems of dimension 1. Imposing periodic boundary conditions, we are thus left with the following boundary value problem:

$$w^{(2k)}(t) = \frac{(-1)^k\mu}{\lambda}w(t) \quad w^{(j)}(0) = w^{(j)}(1) \text{ for } 0 \leq j \leq 2k-1 \tag{3.2}$$

The case of odd j is very similar, in fact $Q_{2k-1}(v)$ can be rewritten as:

$$\begin{aligned}\mathcal{Q}_{2k-1}(v) - \lambda||v||^2 &= \int_0^1 \langle v_{k-1}(t), A_{2k-1}v_k(t)\rangle - \lambda\langle v_0(t), v_0(t)\rangle dt \\ &= \int_0^1 \langle v_0(t), (-1)^{k-1}A_{2k-1}v_{2k-1}(t) - \lambda v_0\rangle dt + \;\; b.t.\end{aligned}$$

Here by *b.t.* we mean boundary terms as the one appearing in the previous equation. They again disappear if we assume that $v_j \in V_j$. Thus we end up with a boundary value problem similar to the one we had before with the difference that now the matrix A_{2k-1} is skew-symmetric.

$$w^{(2k-1)}(t) = \frac{(-1)^{k-1}}{\lambda}A_{2k-1}w(t)$$

If we split the space into the kernel and invariant subspaces on which A_{2k-1} is non degenerate we can decompose $\mathcal{Q}_{2k-1}$ as a direct sum of two-dimensional forms. Imposing periodic boundary conditions, we end up with the following boundary value problems:

$$\begin{cases} w_1^{(2k-1)}(t) &= -\frac{(-1)^{(k-1)}\mu}{\lambda}w_2 \\ w_2^{(2k-1)}(t) &= \frac{(-1)^{(k-1)}\mu}{\lambda}w_1 \end{cases} \quad \begin{cases} w_1^{(j)}(0) = w_1^{(j)}(1), \\ w_2^{(j)}(0) = w_2^{(j)}(1) \end{cases} \quad \text{for } 0 \leq j \leq 2k-2. \tag{3.3}$$

Lemma 3.2. *The boundary value problem in eq. (3.2) has a solution if and only if*

$$\lambda \in \left\{\frac{\mu}{(2\pi r)^{2k}} : r \in \mathbb{N}\right\}.$$

Moreover any such λ has multiplicity 2. In particular, the decreasing sequence of λ for which eq. (3.2) has solutions satisfies:

$$\lambda_r = \frac{\mu}{(2\pi\lceil r/2\rceil)^{2k}} = \frac{\mu}{(\pi r)^{2k}} + O(r^{-(2k+1)}), \quad r \in \mathbb{N}$$

Similarly the boundary value problem in (3.3) has a solution if and only if:

$$\lambda \in \left\{\frac{|\mu|}{(2\pi r)^{2k-1}} : r \in \mathbb{Z}\right\}$$

and any such λ has again multiplicity 2. The monotone rearrangement of λ for which there exists a solution to the boundary value problem is:

$$\lambda_r = \frac{|\mu|}{(2\pi\lceil r/2\rceil)^{2k-1}} = \frac{|\mu|}{(\pi r)^{2k-1}} + O(r^{-(2k)}), \quad r \in \mathbb{Z}$$

Proof. Any solution of the equation $w^{(2k)}(t) = \frac{(-1)^k\mu}{\lambda}w(t)$ can be expressed as a combination of trigonometric and hyperbolic functions with the appropriate frequencies.

Without loss of generality we can assume $\mu > 0$, we have to consider two separate cases:

Case 1: k even and $\lambda > 0$ or k odd and $\lambda < 0$

In this case the quantity $(-1)^k\mu\lambda^{-1} > 0$. If we define $a^{2k} = (-1)^k\mu\lambda^{-1} > 0$ for $a > 0$, we have to solve:

$$w^{(2k)}(t) = a^{2k}w(t), \qquad w^{(j)}(0) = w^{(j)}(1),\ 0 \le j < 2k. \tag{3.4}$$

A base for the space of solutions to the *ODE* is then $\{e^{\omega^j at} : \omega = e^{i\pi/k}\}$. For us it will be more convenient to switch to a real representation of the space of solutions. Notice the following symmetry of the even roots of 1, if η is a root of 1 different form $\pm 1, \pm i$ then $\{\eta, \bar{\eta}, -\eta, -\bar{\eta}\}$ are still distinct roots of 1 (this is also a Hamiltonian feature of the problem).

If we write $\eta = \eta_1 + i\eta_2$, this symmetry implies that the space generated by the following exponential functions $\{e^{\eta t}, e^{\bar{\eta}t}, e^{-\eta t}, e^{-\bar{\eta}t}\}$ is the same as the space generated by

$$\{\sin(\eta_2 t)\sinh(\eta_1 t), \sin(\eta_2 t)\cosh(\eta_1 t), \cos(\eta_2 t)\sinh(\eta_1 t), \cos(\eta_2 t)\cosh(\eta_1 t)\}.$$

Let us rescale these functions by a (so that they solve eq. (3.4)) and call their linear span U_η, we then define U_1 to be the span of $\{\sinh(t), \cosh(t)\}$ and $U_i = \{\sin(t), \cos(t)\}$. Note that U_i appears if and only if k is even.

Thus the solution space for our problem is the space $\bigoplus_\eta U_\eta$ where η ranges over the set $E = \{\eta : \Re(\eta) \ge 0, \Im(\eta) \ge 0, \eta^{2k} = 1\}$.

Now we have to impose the boundary conditions. Notice that, if k is even then U_i is made of periodic functions, so they are always solutions. We can look for more on the complement $\bigoplus_{\eta \ne i} U_\eta$. Suppose by contradiction that w is one of such solutions. Write

$w = \sum_\eta w_\eta$ with $w_\eta \in U_\eta$ and let b be the $\sup\{\Re(\eta) : \eta \in E, w_\eta \neq 0\}$. It follows that either $\sinh(b\,at)$ or $\cosh(b\,at)$ is present in the decomposition of w. It follows that:

$$w(t) = \sinh(b\,at)\frac{w(t)}{\sinh(b\,at)} = \sinh(b\,at)g(t), \quad 0 \not\equiv |g(t)| < C \text{ for } t \text{ large enough}$$

and so $|w|$ is unbounded as $t \to +\infty$ (or $-\infty$) and thus w is not periodic. It follows that there are periodic solutions only if k is even (and thus $\lambda > 0$) and $a = 2\pi r = \sqrt[2k]{\frac{\mu}{\lambda}}$. Notice that we have two independent solutions, so if we order the solution decreasingly we have:

$$\lambda_r = \frac{\mu}{(2\pi\lceil r/2\rceil)^{2k}}, \quad r \in \mathbb{N}$$

Case 2: k odd and $\lambda > 0$ or k even and $\lambda < 0$

In this case we have to look at the roots of -1 but the argument is very similar. If k is even there are no solutions, since you lack purely imaginary frequencies. If k is odd, set $|\mu\lambda^{-1}| = a^{2k}$, then the boundary value problem is:

$$w^{(2k)}(t) = -a^{2k}w(t) \qquad w^{(j)}(0) = w^{(j)}(1),\, 0 \le j < 2k.$$

The roots of -1 are just the roots of 1 rotated by i. Now the space of solutions is $\bigoplus_{\eta\neq 1} U_\eta$. We find again two independent solutions, if we order them we get:

$$\lambda_r = \frac{\mu}{(2\pi\lceil r/2\rceil)^{2k}}, \quad r \in \mathbb{N}$$

Notice that positive μ gives rise to positive solutions. Thus if we consider $\mu < 0$, we get the same result but with switched signs.

We can reduce the odd case (eq. (3.3)) to the even one. Consider the $1-$dimensional equation of twice the order, i.e.:

$$w_1^{2(2k-1)}(t) = -\frac{\mu^2}{\lambda^2}w_1$$

Now, the discussion above tells us that there are exactly two independent solutions with periodic boundary conditions whenever λ satisfies $\sqrt[2k-1]{\frac{\mu}{|\lambda|}} = 2r\pi$. It follows that again there are two independent solutions, this times for both signs of λ. If we order them we get:

$$\lambda_r = \frac{\mu}{(2\pi\lceil r/2\rceil)^{2k-1}}, \quad \lambda_{-r} = \frac{\mu}{(2\pi\lfloor r/2\rfloor)^{2k-1}}, \quad r \in \mathbb{N}$$

□

Proposition 3.1. *Let $\mu > 0$ and $s \in (0, +\infty)$, denote by η_s the number of solutions of eq. (3.2) with λ greater than s and similarly denote by ω_s be the number of solutions with λ bigger than s of:*

$$w^{(2k)}(t) = \frac{(-1)^k\mu}{\lambda}w(t), \quad w^{(j)}(0) = w^{(j)}(1) = 0, \quad k \le j \le 2k-1 \tag{3.5}$$

Then $|\omega_s - \eta_s| \le 2k$. The same conclusion holds for eq. (3.3).

Proof. The result follows from standard results about Maslov index of a path in the Lagrange Grassmannian. References on the topic can be found in [10, 9, 6] and more details are given in Appendix A.2. Let us illustrate briefly the construction. Let (Σ, σ) be a symplectic space, the Lagrange Grassmannian is the collection of Lagrangian subspaces of Σ and it has a structure of smooth manifold. For any Lagrangian subspace L_0 we define the *train* of L_0 to be the set: $T_{L_0} = \{L \text{ Lagrangian} : L \cap L_0 \neq (0)\}$. T_{L_0} is a stratified set, the biggest stratum has codimension 1 and is endowed with a co-orientation. If γ is a smooth curve with values in the Lagrangian Grassmannian (i.e. a smooth family of Lagrangian subspaces) which intersects transversally T_{L_0} in its smooth part, one defines an intersection number by counting the intersection points weighted with a plus or minus sign depending on the co-orientation. Tangent vectors at a point L of the Lagrange Grassmannian (which is a subspace of Σ) are naturally interpreted as quadratic forms on L. We say that a curve is *monotone* if at any point its velocity is either a non negative or a non positive quadratic form. For monotone curves, Maslov index counts the number of intersections with the train up to sign. For generic continuous curves it is defined via a homotopy argument.

Denote by $\mathrm{Mi}_{L_0}(\gamma)$ the Maslov index of a curve γ and L_1 be another Lagrangian subspace. In [6] the following inequality is proved:

$$|\mathrm{Mi}_{L_0}(\gamma) - \mathrm{Mi}_{L_1}(\gamma)| \leq \frac{\dim(\Sigma)}{2} \tag{3.6}$$

Let us apply this results to our problem. First of all let us produce a curve in the Lagrange Grassmannian whose Maslov index coincides with the counting functions ω_s and η_s. The right candidate is the graph of the fundamental solution of $w^{(2k)}(t) = \frac{(-1)^k \mu}{\lambda} w(t)$.

We write down a first order system on $\mathbb{R}^{2k}$ equivalent to our boundary value problem, if we call the coordinates on $\mathbb{R}^{2k}$ x_j, set:

$$x_{j+1}(t) = w^{(j)}(t) \Rightarrow \dot{x}_j = x_{j+1} \text{ for } 1 \leq j \leq 2k-1, \quad \dot{x}_{2k} = \frac{(-1)^k \mu}{\lambda} x_1.$$

For simplicity call $\frac{(-1)^k \mu}{\lambda} = a$, the matrix we obtain has the following structure:

$$A_\lambda = \begin{pmatrix} 0 & & & a \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$$

This matrix is not Hamiltonian with respect to the standard symplectic form on $\mathbb{R}^{2k}$ but is straightforward to compute a similarity transformation that sends it to an Hamiltonian one (recall that we already used that A_λ has the spectrum of an Hamiltonian matrix). Moreover the change of coordinates can be chosen to be block diagonal and thus preserves the subspace $B = \{x_j = 0, k \leq j\}$, which remains Lagrangian too. Since later on we will have to show that the curve we consider is monotone we will give this change of coordinates explicitly. Define the matrix S setting $S_{i,k-i+1} = (-1)^{i-1}$ and zero otherwise. It is a matrix that has alternating ± 1 on the anti-diagonal. Define the following $2k \times 2k$ matrices:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} \quad G^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & (-1)^k S \end{pmatrix} \quad \hat{A}_\lambda = G A_\lambda G^{-1}$$

Set N to be the lower triangular $k \times k$ shift matrix (i.e. the left upper block of A_λ above) and E the matrix with just a 1 in position $(1, k)$ (i.e. the left lower block of A_λ). The new matrix of coefficients is:

$$\hat{A}_\lambda = \begin{pmatrix} N & a(-1)^k ES \\ SE & -N^* \end{pmatrix} \quad ES = \mathrm{diag}(0, \dots, 0, 1), \quad SE = \mathrm{diag}(1, 0, \dots, 0).$$

Now we are ready to define our curve. First of all the symplectic space we are going to use is $(\mathbb{R}^{4k}, \sigma \oplus (-\sigma))$ where σ is the standard symplectic form, in this way graphs of symplectic transformation are Lagrangian subspaces. Sometimes we will denote the direct sum of the two symplectic forms with opposite signs with $\sigma \ominus \sigma$ too. Let Φ_λ be the fundamental solution of $\dot{\Phi}_\lambda^t = \hat{A}_\lambda \Phi_\lambda^t$ at time $t = 1$. Consider its graph:

$$\gamma : \lambda \mapsto \Gamma(\Phi_\lambda^1) = \Gamma(\Phi_\lambda), \quad \lambda \in (0, +\infty)$$

Once we prove that γ is monotone, is straightforward to check that $\mathrm{Mi}_{B\times B}(\gamma|_{[s,+\infty)})$ counts the number of solutions to boundary value problem given in eq. (3.5) for $\lambda \geq s$ and similarly $\mathrm{Mi}_{\Gamma(I)}(\gamma|_{[s,+\infty)})$ counts the solutions of eq. (3.2) for $\lambda \geq s$. Here $\Gamma(I)$ stands for the graph of the identity map (i.e. the diagonal subspace).

Let us check that the curve is monotone. As already mentioned, tangent vectors in the Lagrange Grassmannian can be interpreted as quadratic forms. Being monotone means that the following quadratic form is either non negative or non positive:

$$(\partial_\lambda \gamma)(\xi) = \sigma(\Phi_\lambda \xi, \partial_\lambda \Phi_\lambda \xi), \quad \xi \in \mathbb{R}^{2k}$$

We use the ODE for $\Phi_\lambda(t)$ to prove monotonicity:

$$\begin{aligned} \sigma(\Phi_\lambda \xi, \partial_\lambda \Phi_\lambda \xi) &= \int_0^1 \frac{d}{dt}\big(\sigma(\Phi_\lambda^t \xi, \partial_\lambda \Phi_\lambda^t \xi)\big) dt + \sigma(\Phi_\lambda^0 \xi, \partial_\lambda \Phi_\lambda^0 \xi) \\ &= \int_0^1 \sigma(\hat{A}_\lambda \Phi_\lambda^t \xi, \partial_\lambda \Phi_\lambda^t \xi) + \sigma(\Phi_\lambda^t \xi, (\partial_\lambda \hat{A}_\lambda\, \Phi_\lambda^t + \hat{A}_\lambda \partial_\lambda \Phi_\lambda^t)\xi) dt \\ &= \int_0^1 \sigma(\Phi_\lambda^t \xi, \partial_\lambda \hat{A}_\lambda\, \Phi_\lambda^t \xi) dt \end{aligned}$$

Where we used the facts that $\partial_\lambda \Phi_\lambda^0 = \partial_\lambda Id = 0$ and that $\hat{A}_\lambda$ is Hamiltonian and thus $J\hat{A}_\lambda = -\hat{A}_\lambda^* J$ to cancel the first and third term. It remains to check $J\partial_\lambda \hat{A}_\lambda$. It is straightforward to see that it is a diagonal matrix with just a non zero entry, thus is either non negative or non positive. So $\partial_\lambda \gamma$ is either non positive or non negative being the integral of a non positive or non negative quantity (the sign is independent of ξ).

Now the statement follows from inequality (3.6). □

We are finally ready to compute the asymptotic for $\mathcal{Q}_j$ when the matrix A_j is constant. The next Proposition translate the estimate on the counting functions η_s and ω_s defined in Proposition 3.1 to an estimate for the eigenvalues.

Proposition 3.2. *Let $\mathcal{Q}_j$ be any of the forms appearing in eq. (3.1).*

i) Suppose $j = 2k$ and $\mathcal{Q}_{2k}(v) = \int_0^1 \langle A_{2k} v_k, v_k \rangle dt$ with A_{2k} symmetric and constant and let Σ_{2k} be its spectrum. Define

$$\xi_+ = \left(\sum_{\mu \in \Sigma_{2k}, \mu > 0} \sqrt[j]{\mu} \right)^j \text{ and } \xi_- = \left(\sum_{\mu \in \Sigma_{2k}, \mu < 0} \sqrt[j]{|\mu|} \right)^j.$$

Then $\mathcal{Q}_{2k}$ has capacity (ξ_+, ξ_-) with remainder of order one. Moreover, if A_{2k} is $m \times m$ and $r \in \mathbb{N}$, for $r \geq mk$

$$\frac{\xi_+}{\pi^j(r - 2mk - p(r))^j} \geq \lambda_r \geq \frac{\xi_+}{\pi^j(r + 2mk + p(r))^j} \tag{3.7}$$

where $p(r) = 0$ if r is even or $p(r) = 1$ if r is odd. Similarly for negative r with ξ_-.

ii) Suppose $j = 2k+1$ and $\mathcal{Q}_{2k+1}(v) = \int_0^1 \langle A_{2k+1} v_{k-1}, v_k \rangle dt$ with A_{2k+1} skew-symmetric and constant and let Σ_{2k+1} be its spectrum. Define

$$\xi = \left(\sum_{\mu \in \Sigma_{2k+1}, -i\mu > 0} \sqrt[j]{-i\mu} \right)^j.$$

Then $\mathcal{Q}_{2k+1}$ has capacity ξ with remainder of order one. Moreover , if A_{2k} is $m \times m$ and $r \in \mathbb{Z}$, for $|r| \geq mk$

$$\frac{\xi}{\pi^j(r - 2mk - p(r))^j} \geq \lambda_r \geq \frac{\xi}{\pi^j(r + 2mk + p(r))^j}. \tag{3.8}$$

Proof. First of all we consider 1−dimensional system and we write the inequality $|\eta_s - \omega_s|$ as an inequality for the eigenvalues. Notice that if we have two integer valued function $f, g : \mathbb{R} \to \mathbb{N}$ and an inequality of the form:

$$g(s) \geq \#\{\lambda \text{ solutions of eq. (3.5) } : \lambda \geq s\} \geq f(s),$$

it means that we have at least $f(s)$ solutions bigger than s and at most $g(s)$. This implies that the sequence of ordered eigenvalues satisfies:

$$\lambda_{f(s)} \geq s, \quad \lambda_{g(s)} \leq s.$$

Now we compute this quantities explicitly. In virtue of Proposition 3.1 we can take as upper/lower bounds for the counting function $g(s) = \eta_s + 2k$ and $f(s) = \eta_s - 2k$. We choose the point $s = \frac{\mu}{(2\pi r)^j}$. It is straightforward to see that:

$$\eta_s \Big|_{s = \frac{\mu}{(2\pi r)^j}} = 2\#\{l \in \mathbb{N} : \frac{\mu}{(2\pi l)^j} \geq \frac{\mu}{(2\pi r)^j}\} = 2r.$$

And thus we obtain:

$$\lambda_{2(r-k)} \geq \frac{\mu}{(2\pi r)^j}, \quad \lambda_{2(r+k)} \leq \frac{\mu}{(2\pi r)^j}.$$

Now if we change the labelling we find that , for $l \geq k$:

$$\frac{\mu}{(2\pi(l-k))^j} \geq \lambda_{2l} \geq \frac{\mu}{(2\pi(l+k))^j}.$$

By definition $\lambda_{2l} \geq \lambda_{2l+1} \geq \lambda_{2l+2}$ and thus we have a bound for any index $r \in \mathbb{N}$.

Now we consider m−dimensional system, notice that we reduced the problem, via diagonalization, to the sum of m 1−dimensional systems. Thus our form $\mathcal{Q}_j$ is always a direct sum of 1− dimensional objects. We show now how to recover the desired estimate for the sum of quadratic forms.

First of all observe that counting functions are additive with respect to direct sum. In fact, if $\mathcal{Q} = \oplus_{i=1}^m \mathcal{Q}_i$, λ is an eigenvalue of $\mathcal{Q}$ if and only if it is an eigenvalue of $\mathcal{Q}_i$ for some i. We proceed as we did before. Suppose that $\mathcal{Q}_a$ is $1-$dimensional and $\mathcal{Q}_a(v) = \int_0^1 \mu_a |v_k(t)|^2 dt$. Let us compute η_s in the point $s_0 = (\sum_{i=1}^m \sqrt[j]{\mu_i})^j/(2\pi l)^j$:

$$2\#\left\{r \in \mathbb{N} : \frac{\mu_a}{(2\pi r)^j} \geq \frac{(\sum_{i=1}^m \sqrt[j]{\mu_i})^j}{(2\pi l)^j}\right\} = 2\#\left\{r \in \mathbb{N} : \frac{\sqrt[j]{\mu_a}}{(\sum_{i=1}^m \sqrt[j]{\mu_i})r} \geq \frac{1}{l}\right\}$$

Set for simplicity $c_a = \frac{\sqrt[j]{\mu_a}}{(\sum_{i=1}^m \sqrt[j]{\mu_i})}$, it is straightforward to see that the cardinality of the above set is $\#\{r \in \mathbb{N} : r \leq c_a l\} = \lfloor c_a l \rfloor$. Now we are ready to prove the estimates for the direct sum of forms. Adding everything we have:

$$2\sum_{a=1}^m (\lfloor c_a l \rfloor + k) \geq \#\left\{\text{eigenvalues of } Q \geq \frac{(\sum_{i=1}^m \sqrt[j]{\mu_i})^j}{(2\pi l)^j}\right\} = 2\sum_{a=1}^m (\lfloor c_a l \rfloor - k)$$

It is clear that $\sum_{a=1}^m c_a = 1$ and that $l + mk \geq \sum_{a=1}^m (\lfloor c_a l \rfloor + k)$, similarly $\sum_{a=1}^m (\lfloor c_a l \rfloor + k) \geq l - m(k+1)$ since $\lfloor c_a l \rfloor \geq c_a l - 1$. Rewriting for the eigenvalues with $l \geq mk$ we obtain:

$$\frac{(\sum_{i=1}^m \sqrt[j]{\mu_i})^j}{(2\pi(l - mk))^j} \geq \lambda_{2l} \geq \frac{(\sum_{i=1}^m \sqrt[j]{\mu_i})^j}{(2\pi(l + mk))^j}.$$

It is straightforward to compute the bounds in eqs. (3.7) and (3.8) observing again $\lambda_{2l} \geq \lambda_{2l+1} \geq \lambda_{2l+2}$. □

Remark 3.1. The shift m appearing in eqs. (3.7) and (3.8) is due to the fact we are considering the direct sum of m quadratic forms. It is worth noticing that this does not depend on the fact that we are considering a quadratic form on $L^2([0,1], \mathbb{R}^k)$ and the estimates in eqs. (3.7) and (3.8) hold whenever we consider the direct sum of m $1-$dimensional forms with constant coefficients. This consideration will be used in the proof of Theorem 1.5 below.

3.3 Properties of the capacity

Now we prove some properties of the capacities which are closely related to the explicit estimate we have just proved for the linear case. As done so far we state the proposition for ordered positive eigenvalues. An analogous statement is true for the negative ones.

Proposition 3.3. *Suppose that $\mathcal{Q}$ is a quadratic form on an Hilbert space and let $\{\lambda_n\}_{n\in\mathbb{N}}$ be its positive ordered eigenvalues. Suppose that:*

$$\lambda_n = \frac{\zeta}{n^j} + O(n^{-j-\nu}) \quad \nu > 0, j \in \mathbb{N} \text{ as } n \to +\infty.$$

i) Then for any such $\mathcal{Q}_i$ on a Hilbert space $\mathcal{H}_i$ the direct sum $\mathcal{Q} = \oplus_{i=1}^m \mathcal{Q}_i$ satisfies:

$$\lambda_n = \left(\sum_{i=1}^m \frac{\sqrt[j]{\zeta_i}}{n}\right)^j + O(n^{-j-\nu}) \quad \nu > 0, j \in \mathbb{N} \text{ as } n \to +\infty.$$

ii) Suppose that U is a subspace of codimension $d < \infty$ then

$$\lambda_n(\mathcal{Q}|_U) = \frac{\zeta}{n^j} + O(n^{-j-\nu}) \iff \lambda_n(\mathcal{Q}) = \frac{\zeta}{n^j} + O(n^{-j-\nu}),$$

as $n \to +\infty$.

iii) Suppose that $\mathcal{Q}$ and $\hat{\mathcal{Q}}$ are two quadratic forms. Suppose that $\mathcal{Q}$ is as at the beginning of the proposition and $\hat{\mathcal{Q}}$ satisfies:

$$\lambda_n(\hat{\mathcal{Q}}) = O(n^{j+\mu}) \quad \mu > 0, \text{ as } n \to +\infty.$$

Then the sum $\mathcal{Q}' = \mathcal{Q} + \hat{\mathcal{Q}}$ satisfies:

$$\lambda_n(\mathcal{Q}') = \frac{\zeta}{n^j} + O(n^{j+\nu'}), \qquad \nu' = \min\{\frac{j+\mu}{j+\mu+1}(j+1), j+\nu\}.$$

Proof. The asymptotic relation can be written in terms of a counting function. Take the $j-$th root of the eigenvalues of $\mathcal{Q}_i$, then it holds that

$$\#\{n \in \mathbb{N} \,|\, 0 \le \frac{1}{\sqrt[j]{\lambda_n}} \le k\} = \sqrt[j]{\zeta_i} k + O(k^{1-\nu})$$

So summing up all the contribution we get the estimate in i).

The min-max principle implies that we can control the $n-$th eigenvalue of $\mathcal{Q}|_U$ with the $n-$th and $(n+d)$-th eigenvalue of $\mathcal{Q}$ i.e.:

$$\lambda_n(\mathcal{Q}|_U) \le \lambda_n(\mathcal{Q}) \le \lambda_{n-d}(\mathcal{Q}|_U) \le \lambda_{n-d}(\mathcal{Q})$$

So, if the codimension is fixed, it is equivalent to provide and estimate for the eigenvalues $\mathcal{Q}$ or for those of $\mathcal{Q}|_U$.

For the last point we use Weyl law. We can estimate the $i+j$-th eigenvalue of a sum of quadratic forms with the sum of the $i-$th and the j-th eigenvalues of the summands. Write, as in [7], $\mathcal{Q}'$ as $\mathcal{Q}+\hat{\mathcal{Q}}$ and $\mathcal{Q}$ as $\mathcal{Q}'+(-\hat{\mathcal{Q}})$. and choose $i = n - \lfloor n^\delta \rfloor$ and $j = \lfloor n^\delta \rfloor$ in the first case and $i = n$ and $j = \lfloor n^\delta \rfloor$ in the second. This implies:

$$\lambda_{n+\lfloor n^\delta \rfloor}(\mathcal{Q}) + \lambda_{\lfloor n^\delta \rfloor}(\hat{\mathcal{Q}}) \le \lambda_n(\mathcal{Q}') \le \lambda_{n-\lfloor n^\delta \rfloor}(\mathcal{Q}) + \lambda_{\lfloor n^\delta \rfloor}(\hat{\mathcal{Q}})$$

The best remainder is computed as $\nu' = \max_{\delta\in(0,1)} \min\{(j+\mu)\delta, j+1-\delta, j+\nu\}$. □

Collecting all the facts above we have the following estimate on the decaying of the eigenvalues of $\mathcal{Q}_j$, independently of any analyticity assumption of the kernel.

Proposition 3.4. *Take $\mathcal{Q}_j$ as in the decomposition of lemma (3.1). Then the eigenvalues of $\mathcal{Q}_j$ satisfy:*

$$\lambda_n(\mathcal{Q}_j) = O\Big(\frac{1}{n^j}\Big) \qquad \text{as } n \to \pm\infty$$

Moreover for any $k \in \mathbb{N}$ and for any $0 \le s \le k$ the forms $\mathcal{Q}_{2k+1}$ and $\mathcal{Q}_{2k}$ have the same first term asymptotic as the forms:

$$\hat{\mathcal{Q}}_{2k+1,s}(v) = (-1)^s \int_0^1 \langle A_{2k+1} v_{k+1+s}(t), v_{k-s}(t) \rangle dt$$
$$\hat{\mathcal{Q}}_{2k,s}(v) = (-1)^s \int_0^1 \langle A_{2k} v_{k+s}(t), v_{k-s}(t) \rangle dt$$

Proof. Let's start with even case, $j = 2k$. It holds that:

$$|\mathcal{Q}_{2k}(v)| = |\int_0^1 \langle A_t v_k(t), v_k(t) dt| \leq C \int_0^1 \langle v_k(t), v_k(t)\rangle dt$$

Where $C = \max_t ||A_t||$. By comparison with the constant coefficient case we get the bound.

Suppose now that $j = 2k-1$. As before there is a constant C such that

$$|\mathcal{Q}_{2k}(v)| = |\int_0^1 \langle A_t v_k(t), v_{k+1}(t) dt| \leq C\|v_k\|_2\|v_{k+1}\|_2$$

Consider now the following quadratic forms on $L^2([0,1],\mathbb{R}^k)$:

$$F_k(v) = \int_0^1 ||v_k(t)||^2 dt = \|v_k\|_2^2, \quad F_{k+1}(v) = \int_0^1 ||v_{k+1}(t)||^2 dt = \|v_{k+1}\|_2^2$$

Define $V_n = \{v_1, \dots, v_n\}^\perp$ where v_i are linearly independent eigenvectors of F_k associated to the first n eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$. Similarly define $U_n = \{u_1, \dots, u_n\}^\perp$ to be the orthogonal complement to the eigenspace associated to the first n eigenvalues of F_{k+1}. It follows that:

$$\lambda_{2n}(Q_{2k+1}) \leq \max_{v \in V_n \cap U_n} C\|v_k\|_2\|v_{k+1}\|_2 \leq C \max_{v\in V_n} \|v_k\|_2 \max_{v \in U_n} \|v_{k+1}\|_2$$

We already have an estimate for the eigenvalues of F_k and F_{k+1} since we have already dealt with constant coefficients case. In virtue of the choice of the subspace V_n and U_n, the maxima in the right hand side are the square roots of the $n-th$ eigenvalues of the respective forms. Thus one gives a contribution of order n^{-k} and the other of order n^{-k-1} and the first part of the proposition is proved.

For the second part, without loss of generality suppose that $j = 2k$. The other case is completely analogous.

$$\begin{aligned}\mathcal{Q}_{2k}(v) &= \int_0^1 \langle v_k, A_t v_k\rangle dt = \int_0^1 \langle v_k, \int_0^t A_\tau v_{k-1}(\tau) + \dot{A}_\tau v_k(\tau) d\tau\rangle dt \\ &= -\int_0^1 \langle v_{k+1}(t), A_t v_{k-1}(t) + \int_0^1 \langle v_{k+1}(t), \dot{A}_t v_k(t)\rangle dt\end{aligned}$$

The second term above is of higher order by the first part of the lemma and so iterating the integration by parts on the first term at step s we get that:

$$\begin{aligned}\int_0^1 \langle v_{k+s}(t), A_t v_{k-s}(t)\rangle dt = &-\int_0^1 \langle v_{k+s+1}(t), A_t v_{k-s-1}(t)\rangle dt + \\ &+ \int_0^1 \langle v_{k+s+1}(t), \dot{A}_\tau v_{k-s}(t)\rangle dt\end{aligned}$$

The second term of the right hand side is again of order n^{2k+1}, this can be checked in the same way as in the first part of the proposition. This finishes the proof. □

3.4 Proof of Theorem 1.5

Now we prove the main result of this chapter:

Proof of Theorem 1.5. Suppose that $j = 2k$ is even. We work on $V_k = \{v \in L^2([0,1],\mathbb{R}^m) : v_j(0) = v_j(1) = 0,\ 0 < j \leq k\}$. Then

$$\mathcal{Q}(v) = \mathcal{Q}_{2k}(v) + R_k(v) = \int_0^1 \langle A_t v_k(t), v_k(t)\rangle dt + R_k(v)$$

Since the matrix A_t is analytic we can diagonalize it piecewise analytically in t (see [44]). Thus there exists a piecewise analytic orthogonal matrix O_t such that $O_t^* A_t O_t$ is diagonal. By the second part of Proposition 3.4, if we make the change of coordinates $v_t \mapsto O_t v_t$ we can reduce to study the direct sum of m $1-$ dimensional forms. Without loss of generality we consider forms of the type:

$$\mathcal{Q}_{2k}(v) = \int_0^1 a_t ||v_k(t)||^2 dt = \int_0^1 a_t v_k(t)^2 dt$$

where now a_t is piecewise analytic and v_k a scalar function.

For simplicity we can assume that a_t does not change sign and is analytic on the whole interval. If that were not the case, we could just divide $[0,1]$ in a finite number of intervals and study $\mathcal{Q}_{2k}$ separately on each of them.

Suppose you pick a point t_0 in $(0,1)$ and consider the following subspace of codimension mk in V_k:

$$V_k \supset V_k^{t_0} = \{v \in V_k : v_j(0) = v_j(t_0) = v_j(1) = 0,\ 0 < j \leq k\}$$

For $t \geq t_0$, define $v_j^{t_0} := \int_{t_0}^t v_{j-1}^{t_0}(\tau)d\tau$ and $v_0 = v \in V_k$. It is straightforward to check that on $V_k^{t_0}$ the form Q_{2k} splits as a direct sum:

$$\mathcal{Q}_{2k}(v) = \int_0^{t_0} \langle A_t v_k(t), v_k(t)\rangle dt + \int_{t_0}^1 \langle A_t v_k^{t_0}(t), v_k^{t_0}(t)\rangle dt$$

Now by Proposition 3.3 (points i) and ii)) we can introduce as many points as we want and work separately on each segment and the asymptotic will not change (as long as the number of point is finite).

Now we fix a partition Π of $[0,1]$, $\Pi = \{t_0 = 0, t_1 \dots t_{l-1}, t_l = 1\}$. Consider the subspace $V_\Pi = \{v \in L^2 \,|\, v_s(t_i) = v_s(t_{i+1}) = 0, 0 < s \leq k,\ t_i \in \Pi\}$ which has codimension equal to $k|\Pi|$. Set $a_i^- = \min_{t\in[t_i,t_{i+1}]} a_t$ and $a_i^+ = \max_{t\in[t_i,t_{i+1}]} a_t$. Finally define $v_k^{t_i}(t) = \int_{t_i}^t \cdots \int_{t_i}^{\tau_1} v(\tau)d\tau \dots d\tau_{k-1}$. It follows immediately that on V_Π:

$$\sum_i a_i^- \int_{t_i}^{t_{i+1}} v_k^{t_i}(t)^2 dt \leq \mathcal{Q}_{2k}(v) \leq \sum_i a_i^+ \int_{t_i}^{t_{i+1}} v_k^{t_i}(t)^2 dt$$

Now, we already analysed the spectrum for the problem with constant a_t on $[0,1]$. The last step to understand the quantities on the right and left hand side is to see how the eigenvalues rescale when we change the length of $[0,1]$.

If we look back at the proof of Lemma 3.2, it is straightforward to check that the length is relevant only when we impose the boundary conditions, we find that the eigenvalues are: $\lambda = \frac{a\ell^{2k}}{(2\pi n)^{2k}}$ and again double.

In particular the estimates in eqs. (3.7) and (3.8) are still true replacing μ_i with $a_i^{\pm}\ell^{2k}$.
If we replace now ℓ by $|t_{i+1}-t_i|$ and sum the capacities according to Proposition 3.3 we have the following estimate on the eigenvalues on V_Π, for $n \geq 2k|\Pi|$:

$$\Big(\frac{\sum_i (a_i^-)^{\frac{1}{2k}}(t_{i+1}-t_i)}{\pi(n+2|\Pi|k+p(n))}\Big)^{2k} \leq \lambda_n(\mathcal{Q}_{2k}\big|_{V_\Pi}) \leq \Big(\frac{\sum_i (a_i^+)^{\frac{1}{2k}}(t_{i+1}-t_i)}{\pi(n-2|\Pi|k-p(n))}\Big)^{2k}$$

Moreover the min-max principle implies that, for $n \geq k|\Pi|$:

$$\lambda_n(\mathcal{Q}_{2k}\big|_{V_\Pi}) \leq \lambda_n(\mathcal{Q}_{2k}) \leq \lambda_{n-k|\Pi|}(\mathcal{Q}_{2k}\big|_{V_\Pi})$$

In particular for $n \geq 3k|\Pi|$ we have:

$$\Big(\frac{\sum_i (a_i^-)^{\frac{1}{2k}}(t_{i+1}-t_i)}{\pi(n+2|\Pi|k+p(n))}\Big)^{2k} \leq \lambda_n(Q_{2k}) \leq \Big(\frac{\sum_i (a_i^+)^{\frac{1}{2k}}(t_{i+1}-t_i)}{\pi(n-3|\Pi|k-p(n))}\Big)^{2k} \tag{3.9}$$

We address now the issue of the convergence of the Riemann sums. Set:

$$I_a^{\pm} = \sum_i (a_i^{\pm})^{\frac{1}{2k}}(t_{i+1}-t_i)$$

and $I_a = \int_0^1 a^{\frac{1}{2k}}dt$. It is well know that $I_a^{\pm} \to I_a$ as long as $\sup_i |t_i - t_{i+1}|$ goes to zero. We need a more quantitative bound on the rate of convergence. Using results from [24] for and equispaced partition, we have that:

$$|I_a - I_a^{\pm}| \leq C_a^{\pm}\frac{1}{|\Pi|} = \frac{C(a,k,\pm)}{\mathrm{codim}(V_\Pi)}$$

Where $C(a,k,\pm)$ is a constant that depends only on the function a and on k and the inequality holds for $|\Pi| \geq n_0$ sufficiently large, where n_0 depends just on a and k.

Consider the right hand side of eq. (3.9), adding and subtracting $\frac{I_a}{(\pi n)^{2k}}$, we find that for $n \geq \max\{n_0, k|\Pi|\}$:

$$\lambda_n(\mathcal{Q}_{2k}) \leq \Big(\frac{I_a}{\pi n}\Big)^{2k} + \Big(\frac{I_a^+}{\pi(n-3|\Pi|k-p(n))}\Big)^{2k} - \Big(\frac{I_a}{\pi n}\Big)^{2k}.$$

A simple algebraic manipulation shows that there are constants C_1, C_2 and C_3 such that the difference on the right hand side is bounded by

$$\frac{C_1 n^{2k}|\Pi|^{-1} + C_2(n^{2k} - |\Pi|^{2k}(n/|\Pi|-1)^{2k})}{C_3(n-3k|\Pi|)^{2k}n^{2k}}$$

for $n \geq \max\{3k|\Pi|, n_1|\Pi|, n_0\}$ where n_1 is a certain threshold independent of $|\Pi|$.

The idea now is to choose for n a partition Π of size $|\Pi| = \lfloor n^\delta \rfloor$ to provide a good estimate of $\lambda_n(\mathcal{Q})$. The better result in terms of approximation is obtained for $\delta = \frac{1}{2}$. Heuristically this can be explained as follows: on one hand the first piece of the error term is of order $n^{-2k-\delta}$, comes from the convergence of the Riemann sums and gets better as $\delta \to 1$. On the other hand the second term comes from the estimate on the eigenvalues and get worse and worse as n^δ becomes comparable to n.

A perfectly analogous argument allows to construct an error function for the left side of eq. (3.9) which decays as $n^{-2k-1/2}$ for n sufficiently large.

We have proved so far that, if we are dealing with quadratic forms on scalar functions, $\mathcal{Q}_{2k}$ has $2k-$capacity $\xi_+ = (\int_0^1 \sqrt[2k]{a_t}dt)^{2k}$. Now we apply point $i)$ of Proposition 3.3 to obtain the formula in the statement for forms on $L^2([0,1],\mathbb{R}^m)$. Finally notice that by Proposition 3.4 the eigenvalues of $R_k(v)$ decay as n^{-2k-1}. If we apply point $iii)$ of Proposition 3.3 we find that $\mathcal{Q}_{2k}(v)+R_k(v)$ has the same $2k-$capacity as $\mathcal{Q}_{2k}$ with remainder of order $1/2$.

Now we consider the case $j = 2k-1$. The idea is to reduce to the case of $j = 4k-2$ as in the proof of Lemma 3.2 and use the symmetries of $\mathcal{Q}_{2k-1}$ to conclude. In the same spirit as in the beginning of the proof let us diagonalize the kernel A_{2k-1}. We thus reduce everything to the two dimensional case, i.e. to the quadratic forms:

$$\mathcal{Q}(v) = \int_0^1 \langle v_k(t), \begin{pmatrix} 0 & -a_t \\ a_t & 0 \end{pmatrix} v_{k-1}(t)\rangle dt \quad a_t \geq 0 \tag{3.10}$$

It is clear that the map $v_0 \mapsto Ov_0$ where $O = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is an isometry of $L^2([0,1],\mathbb{R}^2)$ and $\mathcal{Q}(Ov_0) = -\mathcal{Q}(v_0)$ and so the spectrum is two sided and the asymptotic is the same for positive and negative eigenvalues.

Now we reduce the problem to the even case. Let's consider *the square* of $\mathcal{Q}_{2k-1}$. By proposition (3.4) $\mathcal{Q}_{2k-1}$ has the same asymptotic as the form:

$$\hat{\mathcal{Q}}_{2k-1} = (-1)^{k+1}\int_0^1 \langle A_t v_{2k-1}(t), v_0(t)\rangle dt \qquad F(v_0)(t) = (-1)^{k+1}A_t v_{2k-1}(t)$$

So we have to study the eigenvalues of the symmetric part of F. It is clear that:

$$\frac{(F+F^*)^2}{4} = \frac{F^2+FF^*+F^*F+(F^*)^2}{4}$$

Thus we have to deal with the quadratic form:

$$\begin{aligned} 4\tilde{\mathcal{Q}}(v) &= \langle [2F^2+F^*F+FF^*](v), v\rangle \\ &= 2\langle F(v), F^*(v)\rangle + \langle F^*(v), F^*(v)\rangle + \langle F(v), F(v)\rangle \end{aligned}$$

The last term is the easiest to write, it is just:

$$\langle F(v), F(v)\rangle = \int_0^1 \langle -A_t^2 v_{2k-1}(t), v_{2k-1}(t)\rangle dt$$

which is precisely of the form of point $i)$ and gives $\frac{1}{4}$ of the desired asymptotic. The operator F^* acts as follows:

$$F^*(v) = (-1)^{k+1}\int_0^t \int_0^{t_{2k-1}} \dots \int_0^{t_1} A_{t_1} v_0(t_1) dt_1 \dots dt_{2k-1}$$

Using integration by parts one can single out the term $A_t v_{2k-1}$. To illustrate the procedure, for $k=1$ one gets:

$$\begin{aligned} F^*(v) &= A_t v_1(t) - \int_0^t \dot{A}_\tau v_1(\tau)d\tau \\ \langle F^*(v), F^*(v)\rangle &= \int_0^1 \langle -A_t^2 v_1(t), v_1(t)\rangle dt + 2\int_0^1 \langle A_t v_1(t), \int_0^t \dot{A}_\tau v_1(\tau)d\tau\rangle dt + \\ &\quad + \int_0^1 \langle \int_0^t \dot{A}_\tau v_1(\tau)d\tau, \int_0^t \dot{A}_\tau v_1(\tau)d\tau\rangle dt \end{aligned}$$

The other terms thus do not affect the asymptotic since by Proposition 3.4 they decay at least as $O(n^3)$. The proof goes on the same line for general k.

The same reasoning applies to the term $\langle F(v), F^*(v)\rangle$. Summing everything one gets that the leading term is $\int_0^1 \langle -A_t^2 v_{2k-1}(t), v_{2k-1}(t)\rangle dt$ and so this is precisely the same case as point i). Recall that A_t is a 2×2 skew-symmetric matrix as defined in eq. (3.10), thus the eigenvalues of the square coincide and are a_t^2. It follows that, for n sufficiently large, the square of the eigenvalues of $\tilde{\mathcal{Q}}$ satisfy:

$$\lambda_n(\tilde{\mathcal{Q}}) = \frac{\left(\int_0^1 2\sqrt[4k-2]{a_t^2}dt\right)^{4k-2}}{\pi^{4k-2}n^{4k-2}} + O(n^{-4k-2-\frac{1}{2}})$$

It is immediate to see that $\frac{\left(\int_0^1 2\sqrt[4k-2]{a_t^2}dt\right)^{4k-2}}{(\pi n)^{4k-2}} = \frac{\left(\int_0^1 \sqrt[2k-1]{a_t}dt\right)^{4k-2}}{(\pi n/2)^{4k-2}}$. This mirrors the fact that the spectrum of $\mathcal{Q}_{2k-1}$ is double and any couple $\lambda, -\lambda$ is sent to the same eigenvalue λ^2. Thus the $(2k-1)-$capacity of $\mathcal{Q}_{2k-1}$ is $(\int_0^1 \sqrt[2k-1]{a_t}dt)^{2k-1}$.

Moreover, given two sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, $\sqrt{a_n^2+b_n^2} = a_n\sqrt{1+\frac{b_n^2}{a_n^2}} \approx a_n(1+\frac{b_n}{a_n} + O(\frac{b_n}{a_n}))$ so the remainder is still $2k-1+\frac{1}{2}$.

Arguing again by point i) of Proposition 3.3 one gets the estimate in the statement.

The last part about the $\infty-$capacity follow just by Proposition 3.4. If $A_j \equiv 0$ for any j then for any $\nu\in\mathbb{R}$, $\nu>0$ we have $\lambda_n n^\nu \to 0$ as $n\to\pm\infty$. □

Remark 3.2. We can interpret Theorem 1.5 as a quantitative version of various necessary optimality conditions that one can formulate for certain classes of singular extremals (see [12, Chapter 20] or [8, Chapter 12]). Moreover, leaving optimality conditions aside, Theorem 1.5 gives the asymptotic distribution of the eigenvalues of the second variation for totally singular extremals (see definition B.2).

As mentioned in appendix B we can produce a representation of the second variation also in the non strictly normal case which is at least formally very similar to the normal case. However, a common occurrence is that the matrix H_t completely degenerates and is constantly equal to the zero matrix. This is the case for affine control systems and abnormal extremal in Sub-Riemannian geometry, i.e. systems of the form:

$$f_u = \sum_{i=1}^{l} f_i u_i + f_0, \quad f_i \text{ smooth vector fields}$$

In this case Legendre condition $H_t \le 0$ (see the previous section) does not give much information. One, then, looks for *higher* order optimality conditions. This is usually done exactly as in Lemma 3.1: the first optimality conditions one finds are *Goh condition* and *generalized Legendre condition* which prevent the second variation from being *strongly indefinite*.

In the notation of Lemma 3.1 Goh conditions is written as $\mathcal{Q}_1 \equiv 0$ i.e. $Z_t^* J Z_t \equiv 0$. It can be reformulated in geometric terms as follows, if λ_t is the extremal then

$$\lambda_t[\partial_u f_u(q(t))v_1, \partial_u f_u(q(t))v_2] = 0, \;\forall\, v_1, v_2 \in \mathbb{R}^k$$

From Theorem 1.5 it is clear that if $\mathcal{Q}_1 \not\equiv 0$, the second variation has infinite negative index and that eigenvalues distribute evenly between the negative and positive parts of the spectrum. Then one asks that the second term $\mathcal{Q}_2$ is non positive definite, otherwise the negative part of the spectrum of $-\mathcal{Q}_2$ becomes infinite. In our notation this condition reads

$$(Z_t^{(1)})^* J Z_t \leq 0 \iff \sigma(Z_t^{(1)}v, Z_t v) \leq 0,\, \forall\, v \in \mathbb{R}^k.$$

Again it can be translated in a differential condition along the extremal, however this time it will in general involve more than just commutators if the system is not control affine.

If $\mathcal{Q}_2 \equiv 0$, one can take more derivatives and find new conditions. In particular, using the notation of Lemma 3.1, one has always to ask that the first non zero term in the expansion is of even order and that the matrix of its coefficients is non positive in order to have finite negative index.

Chapter 4

The Morse index of the Second Variation

This chapter focuses on the computation of the Morse index of the Second Variation for problems with moving endpoints. We give a proof of the formula in Theorem 1.1 and apply it to some concrete situations. For example we give iteration formulae for extremals under periodic boundary conditions and *discretization* formulae which reduce the problem of computing the index to a finite sum of finite dimensional contributions.

The results of this chapter are mainly contained in [15] and were obtained jointly with A. Agrachev and I. Beschastnyi. The topic is classical and there are various sources for similar formulae in literature tracing back to [22, 30, 53, 25, 30]. A result similar to Theorem 1.1 has been obtain independently in [41]. Our proofs have a different flavour than the ones involving intersection theory in the Lagrange Grassmannian. They essentially relay on linear algebra and the structure of the Second Variation.

Through out this section we will use the notation introduce in Sections 1.1 and 2.3 and appendices A.2 and B. In particular we will use the construction in Section 2.3 to reduce moving boundary conditions to fixed points one.

Before going to the proof of the main result however, we will focus on some applications. In the upcoming section we will give a formula to compute the Morse index of an extremal substituting the infinite dimensional space of variations with a finite dimensional one. We will prove a *filtration* formula for minimization problems on graphs, in the spirit of Example 1.1 of Section 1.1. Lastly we will prove iteration formulae for periodic extremal giving counterparts of the results in [22, 30].

4.1 Discretization formula

We will work with an optimal control problem as given in eq. (B.2) and a strictly normal extremal λ_t. In order to formulate the next result, we need the definition of conjugate times and conjugate points (for Dirichlet boundary conditions).

Definition 4.1. Given $\mu \in T^*M$, denote by Π_μ the vertical subspace $T_\mu(T^*_{\pi(\mu)}M)$. Given an extremal $\lambda : [0,1] \to T^*M$ of an optimal control problem as in eq. (B.2), let Φ_t be the fundamental solution of *Jacobi* equation (eq. (1.2)) at time t. we say that an instant of time $t \in [0,1]$ is *conjugate* if the map

$$\pi_* \circ (\Phi_t)|_{\Pi_{\lambda(0)}}$$

has a kernel. The corresponding point $q(t) = \pi(\lambda(t))$ is said to be a *conjugate point.*

The above definition is completely analogous to the classical given in terms of the linearisation of the extremal flow (when the latter is defined). In fact the map $\tilde{\Phi}_t$ introduced in eq. (B.11) preserves the submanifolds $\Pi_m = \pi^{-1}(m)$ for any $m \in M$ and the symplectomorphism $(\tilde{\Phi}_t)_*\Phi_t$ coincides with the fundamental solution of the linearisation of the extremal flow (see [12][Proposition 21.3]) . To simplify notation we will denote by $\Theta_t = (\tilde{\Phi}_t)_*\Phi_t$ and $\Theta_{s,t}$ the map $\Theta_t\Theta_s^{-1}$. The latter notation is nothing else than the standard notation for non autonomous linear flows. We will denote by $\Gamma(\Theta_t)$ its graph and by $\Gamma(\Theta)$ the graph at time $t = 1$.

A consequence of Theorem 1.1 is the following result.

Theorem 4.1 (Discretization)**.** *Let $\lambda : [0, 1] \to T^*M$ be an extremal as above with Dirichlet boundary conditions, i.e. $N = \{q_0\} \times \{q_1\}$. Let $\Xi = \{t_0, \dots, t_n\}$ be a partition of $[0, 1]$. The following formula holds:*

$$\operatorname{ind}^- Q \geq \sum_{i=0}^{n-1} i\big(\Theta_{i+1,i}^{-1}(\Pi_{i+1}), \Pi_i, \Theta_{i,i-1} \circ \dots \circ \Theta_{1,0}(\Pi_0)\big), \tag{4.1}$$

where $\Pi_i = T_{\lambda(t_i)}(T^*_{\pi(\lambda(t_i))}M) \simeq T^*_{\pi(\lambda(t_i))}M$. Moreover, equality holds if $\max_i |t_{i+1} - t_i|$ is sufficiently small and no t_i is a conjugate time.

In particular, strong Legendre condition along the extremal, ensures that the Morse index is finite and that the conjugate points form a discrete set. This will guarantee that, under mild conditions and after enough successive refinements of the partition, formula (4.1) will give exactly the Morse index of the extremal. The following picture illustrate how the admissible variation in the case we fix some intermediate points $\{q(t_i)\}$ change.

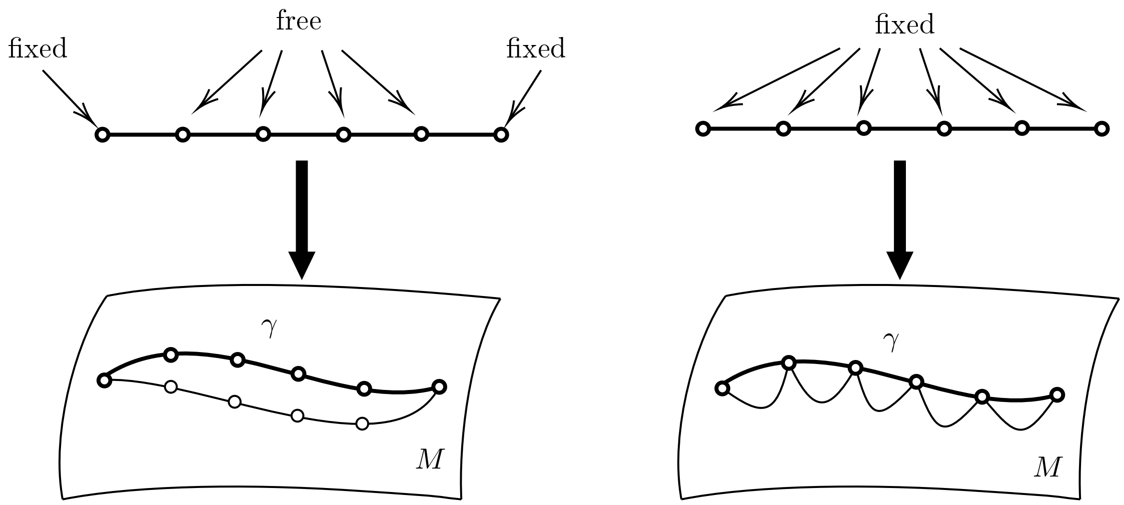


Figure 4.1: Variation of γ in the original problem and a problem with extra fixed vertices.

Let us first prove the formula when only one extra vertex is introduced. Let $\gamma = \pi(\lambda)$ be an extremal curve in a problem with fixed end-points. Take a point $t^* \in (0, 1)$. Let us call $\gamma_1 = \gamma|_{[0,t^*]}$ and $\gamma_2 = \gamma|_{[t^*,1]}$ the restrictions. Q_{γ_i} will denote the second variation of the segment as an extremal curve with fixed points. Recall that $\Pi_i = T_{\lambda(t_i)}(T^*_{\pi(\lambda(t_i))}M) \simeq T^*_{\pi(\lambda(t_i))}M$ is the vertical subspace over the point $\gamma(t_i)$.

Proposition 4.1. *The index of the second variation Q_γ can be computed as follows:*

$$\operatorname{ind}^- Q_\gamma = \operatorname{ind}^- Q_{\gamma_1} + \operatorname{ind}^- Q_{\gamma_2} + i\big(\Theta_2^{-1}(\Pi_2), \Pi_1, \Theta_1(\Pi_0)\big) + k, \tag{4.2}$$

where $k = \dim(\Theta_2(\Pi_1) \cap \Pi_2) + \dim(\Theta_1(\Pi_0) \cap \Pi_1) - \dim(\Theta_2^{-1}(\Pi_2) \cap \Pi_1 \cap \Theta_1(\Pi_0))$.

Proof. Let us consider the following three points in M:

$$q_0 = \gamma(0), \quad q_1 = \gamma(t^*), \quad q_2 = \gamma(1).$$

Variations of γ as a curve form q_0 to q_2 do not necessarily pass by the point q_1 at time t^* but satisfy a continuity condition there. Instead of considering a problem on two segments we double the state space and work on just one interval. To do this we break up $[0,1]$ in two intervals and consider the dynamics separately (i.e. duplicate the variables). The new boundary conditions which allow us to glue the two pieces together are of the form:

$$(\gamma_1(0), \gamma_2(t^*), \gamma_1(t^*), \gamma_2(1)) \in \{q_0\} \times \Delta \times \{q_2\} = \{(q_0, q_1, q_1, q_2) |\, q_1 \in M\}.$$

Now we are going to compare the following two problems. The first one is *fixed end-points*, we impose that the curve starts from (q_0, q_1) and arrives to (q_1, q_2). The second one is curves satisfying the constraints given by the manifold $N = \{q_0\} \times \Delta \times \{q_2\}$ defined above.

Recall that γ is a projection of a solution $\lambda : [0,1] \to T^*M$ of the Hamiltonian system. Let us consider the tangent space to the annihilator of N at the point $\underline{\lambda} = (\lambda(0), \lambda(1), \lambda(1), \lambda(2))$. Fix a system of coordinates, which determines a complement to the subspace $\ker \pi_* = \Pi_1$ which we call B. In these coordinates the annihilator reads:

$$T_{\underline{\lambda}} A(N) = \left\{ \begin{pmatrix} \nu_1 \\ \alpha + X \\ \alpha + X \\ \nu_2 \end{pmatrix} : \alpha, \nu_i \in \Pi_1,\ X \in B \right\}.$$

The other space appearing is the graph of the two symplectomorphisms Θ_1 and Θ_2 coming from the Hamiltonian flows of PMP on intervals $[0, t^*]$ and $[t^*, 1]$. It will be denoted by $\Gamma(\Theta_1 \times \Theta_2)$.

Let us look at the subspace on which the Maslov form m is defined, $(T_{\underline{\lambda}} A(N) + \Pi^4) \cap \Gamma(\Theta_1 \times \Theta_2)$, where $\Pi^4 = \Pi_0 \times \Pi_1^2 \times \Pi_2$. This is defined by the following equations,

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \Theta_1(\xi_1) \\ \Theta_2(\xi_2) \end{pmatrix} = \begin{pmatrix} \nu_1 \\ \alpha + X \\ \alpha + X \\ \nu_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} \iff \xi_1 \in \Pi_0, \Theta_2(\xi_2) \in \Pi_2,\ \xi_2 - \Theta_1(\xi_1) \in \Pi_1,$$

where α, ν_i, for $i = 1, 2$ and μ_j for $j = 1, \dots 4$ lie in the vertical subspace over the respective points, whereas $X \in B$ is in the horizontal space. In particular Maslov form reads:

$$m(\xi_1, \xi_2) = \sigma(\mu_3 - \mu_2, X) = \sigma(\Theta_1(\xi_1) - \xi_2, \xi_2) = \sigma(\Theta_1(\xi_1), \xi_2) = \sigma(\xi_2, -\Theta_1(\xi_1)).$$

So, if we call $\eta = \Theta_2(\xi_2) \in \Pi_2$ and $\xi = \xi_1$ we have $\xi \in \Pi_0, \eta \in \Pi_2$ and $\Theta_2^{-1}(\eta) - \Theta_1(\xi) \in \Pi_1$ and see that the form coincides with:

$$m(\Pi^4, \Gamma(\Theta_1 \times \Theta_2), T_{\underline{\lambda}} A(N)) = m(\Theta_2^{-1}(\Pi_2), \Pi_1, \Theta_1(\Pi_0)).$$

The additional terms popping up in Theorem 1.1 are

$$\dim(\Gamma(\Theta_1 \times \Theta_2) \cap \Pi^4) = \dim(\Theta_1(\Pi_0) \cap \Pi_1) + \dim(\Theta_2^{-1}(\Pi_2) \cap \Pi_1)$$

and

$$\dim(\Gamma(\Theta_1 \times \Theta_2) \cap \Pi^4 \cap T_{\underline{\lambda}} A(N)) = \dim(\Theta_2^{-1}(\Pi_2) \cap \Theta_1(\Pi_0) \cap \Pi_1)$$

as a quick calculation shows. □

We can now prove Theorem 4.1.

Prood of Theorem 4.1. The statement will follow from Proposition 4.1. First of all notice that in eq. (4.2) all terms are positive, this gives easily that $\operatorname{ind}^- Q \geq i\big(\Theta_2^{-1}(\Pi), \Pi, \Theta_1(\Pi)\big)$ when the partition is $\Xi = \{0, t^*, 1\}$. For a general Ξ apply Proposition 4.1 iteratively to $\{0, t_{j-1}, t_j\}$ where j runs from 2 to n. This allows to express the index of the second variation of $\gamma|_{[0,t_j]}$ as the sum of the index of the second variation of $\gamma|_{[0,t_{j-1}]}$ and $\gamma_j := \gamma|_{[t_{j-1},t_j]}$ plus Maslov index terms and dimensions of intersections.

Iteratively replacing the terms $\operatorname{ind}^- Q_{\gamma|_{[0,t_j]}}$ we obtain the following formula:

$$\operatorname{ind}^- Q_\gamma = \sum_{j=0}^{n-1} \operatorname{ind}^- Q_{\gamma_j} + i\big(\Theta_{j+1,j}^{-1}(\Pi_{j+1}), \Pi_j, \Theta_{j,j-1} \circ \cdots \circ \Theta_{1,0}(\Pi_0)\big)$$
$$+ \dim(\Theta_{j,0}(\Pi_0) \cap \Pi_j) + \dim(\Theta_{j+1,j}(\Pi_j) \cap \Pi_{j+1}) - \dim(\Theta_{j,0}(\Pi_0) \cap \Pi_j \cap \Theta_{j+1,j}^{-1}(\Pi_{j+1})).$$

Here the maps $\Theta_{j,j-1}$ are defined as in the statement of Theorem 4.1. The notation is related to law of composition of non autonomous flows and is justified by the fact that $\Theta_{j,k} \circ \Theta_{k,l} = \Theta_{j,l}$.

The index is presented as sum of three positive terms: the first one Q_{γ_i} is zero when each segment $\gamma|_{[t_i,t_{i+1}]}$ is minimizing [12, Theorem 21.3]. Under Legendre strong conditions this is the case when $\sup_i |t_i - t_{i-1}|$ is small enough (see [12] for instance). The same goes for $\dim(\Theta_{i+1,i}(\Pi_i) \cap \Pi_{i+1}) - \dim(\Theta_{i,0}(\Pi_0) \cap \Pi_i \cap \Theta_{i+1,i}^{-1}(\Pi_{i+1}))$. Moreover $\dim(\Theta_{i,0}(\Pi_0) \cap \Pi_i)$ is zero precisely when t_i is not a conjugate time for γ.

Thus equality holds exactly when our hypotheses on the partition are satisfied. □

Remark 4.1. The hypothesis on the partition Ξ can be weakened if we change a bit our way of counting. If we add to the dimension of the negative space the dimension of the null space of the Maslov form we can essentially forget about avoiding conjugate points of γ.

You can see that the correction term k in Proposition 4.1 is in fact the dimension of the kernel of the Maslov form $m\big(\Theta_2^{-1}(\Pi_2), \Pi_1, \Theta_1(\Pi_1)\big)$. The quantity:

$$\sum_{i=1}^{n-1} (\operatorname{ind}^- + \ker)\Big(m\big(\Theta_{i+1,i}^{-1}(\Pi_{i+1}), \Pi_i, \Theta_{i,i-1} \circ \cdots \circ \Theta_{1,0}(\Pi_0)\big)\Big)$$

still approximates from below the negative index and includes the contribution of conjugate points of γ that are possibly present in the partition.

Remark 4.2. If we combine Theorem 1.1 and Theorem 4.1 we can obtain a formula for the index involving just the Maslov index i and dimension of intersections for arbitrary boundary conditions.

4.2 Filtration formula

In the previous subsection we have proven a discretization formula for the fixed end-point problem on an interval. The idea was to introduce extra vertices inside the single edge and apply an iterative procedure of fixing each of the new vertices one by one. Note that if we would have fixed all of the vertices at the same time, a direct application of eq. (1.3) would result in computation of the Maslov index in a symplectic space of a very big dimension. Instead the recursive nature of the proof allows us to reduce greatly the dimensionality of the problem.

A way of reducing the dimensionality in formula (1.3) for problems with separated boundary conditions is discussed in paper [19]. The argument works when all of the Lagrangian spaces in the final formula are transversal. We can reproduce this argument in a greater generality using Theorem 1.1.

Assume that each vertex $v \in \mathcal{G}_0$ is constrained to lie on a separate submanifold $N_v \subset M$. We construct a manifold N of boundary conditions in $M^{2|\mathcal{G}_1|}$ in the following way: define the map $\hat{s} : M^{|\mathcal{G}_1|} \to M^{|\mathcal{G}_0|}$ sending $m_e \to m_{s(e)}$. Here s (and t) denote the *source* and *target* maps as defined in Section 1.1. Similarly we get $\hat{t} : M^{|\mathcal{G}_1|} \to M^{|\mathcal{G}_0|}$ sending $m_e \to m_{t(e)}$. We then define $N = (\hat{s} \times \hat{t})^{-1}(\prod_{v \in \mathcal{G}_0} N_v)$ which now is a subset of $M^{2|\mathcal{G}_1|}$.

We can introduce a filtration of vertices

$$\emptyset = \mathcal{G}_0^0 \subset \mathcal{G}_0^1 \subset \cdots \subset \mathcal{G}_0^{|\mathcal{G}_0|} = \mathcal{G}_0,$$

such that

$$|\mathcal{G}_0^j| = |\mathcal{G}_0^{j-1}| + 1, \qquad i = 1, \dots, |\mathcal{G}_0|.$$

To each set $\mathcal{G}_0^j$ we associate boundary conditions $N_j \subset N$ in the following way. We assume that vertices $v \in \mathcal{G}_0^j$ vary on N_v, while vertices $v \in \mathcal{G}_0 \setminus \mathcal{G}_0^j$ are assumed to be fixed and perform the construction explained few lines above. Thus we activate variations of each individual vertex at a time and track how the index changes as we do so.

We now apply Theorem 1.1 to compute $\operatorname{ind}^- Q_{N_{j+1}} - \operatorname{ind}^- Q_{N_j}$. Let us introduce some simplifying notations. Recall that $s, t : \mathcal{G}_1 \to \mathcal{G}_0$ are the source and the target maps. Let $v_j \in \mathcal{G}_0^{j+1} \setminus \mathcal{G}_0^j$ be the activated vertex. We introduce a separate notation for the set of edges that are incident to v_j:

$$\mathcal{G}_1^j = s^{-1}(v_j) \cup t^{-1}(v_j).$$

A naive guess would be that when we activate a vertex, the only relevant contributions come from the edges incident to a given vertex. Thus we define forgetful projections $\pi_{\mathcal{G}_1^j}$ which forget all the edges except the ones incident to v_j:

$$\pi_{\mathcal{G}_1^j} : T_{\bar{\lambda}}(T^*M)^{\mathcal{G}_1} \times T_{\bar{\lambda}}(T^*M)^{\mathcal{G}_1} \to T_{\bar{\lambda}}(T^*M)^{\mathcal{G}_1^j} \times T_{\bar{\lambda}}(T^*M)^{\mathcal{G}_1^j}.$$

Subspaces $T_{\bar{\lambda}}A(N_{j-1})$ and $T_{\bar{\lambda}}A(N_j)$ can have a big intersection. For sure this intersection contains the subset $V_j = \pi_{\mathcal{G}_1^j}^{-1}(0)$, which is an isotropic subspace. This means that we can perform a symplectic reduction to the space $V_j^{\perp}/V_j$. Let

$$\pi_j : T_{\bar{\lambda}}(T^*M)^{|\mathcal{G}_1|} \times T_{\bar{\lambda}}(T^*M)^{|\mathcal{G}_1|} \to V_j^{\perp}/V_j$$

be the projection maps for each $j = 1, \ldots, |\mathcal{G}_1|$. We can then define shortened notations for the images:

$$\begin{aligned} A^j_{j-1} &= \pi_j(T_{\bar\lambda}A(N_{j-1})), \\ A^j_j &= \pi_j(T_{\bar\lambda}A(N_j)), \\ \Gamma(\Theta_j) &= \pi_j(\Gamma(\Theta)). \end{aligned}$$

By property (A.6), we can factor out V_j in the definition of the Maslov index and get

$$i(T_{\bar\lambda}A(N_{j-1}), \Gamma(\Theta), T_{\bar\lambda}A(N_j)) = i(A^j_{j-1}, \Gamma(\Theta_j), A^j_j)$$

and for the same reason

$$\begin{aligned} \dim\left(T_{\bar\lambda}A(N_{j-1}) \cap \Gamma(\Theta)\right) - \dim\left(T_{\bar\lambda}A(N_j) \cap \Gamma(\Theta) \cap T_{\bar\lambda}A(N_{j-1})\right) = \\ = \dim\left(A^j_{j-1} \cap \Gamma(\Theta_j)\right) - \dim\left(A^j_{j-1} \cap \Gamma(\Theta_j) \cap A^j_j\right). \end{aligned}$$

Finally since $N_{j-1} \subset N_j$, we have that

$$\dim(T_{\pi(\bar\lambda)}N_{j-1} \cap T_{\pi(\bar\lambda)}N_j) - \dim(T_{\pi(\bar\lambda)}N_{j-1}) = 0.$$

Now we collect all of the terms and sum by the index $j = 1, \ldots, |\mathcal{G}_0|$. As the result we obtain a formula that expresses the difference between the index of the second variation Q of the original problem with the index of the second variation $Q_0 := Q_{N_0}$ of the problem with the same graph and fixed vertices:

$$\begin{aligned} \operatorname{ind}^- Q - \operatorname{ind} Q_0 &= \sum_{j=1}^{|\mathcal{G}_0|} \operatorname{ind}^- Q_{N_j} - \operatorname{ind}^- Q_{N_{j-1}} \\ &= \sum_{j=1}^{|\mathcal{G}_0|} i(A^j_{j-1}, \Gamma(\Theta_j), A^j_j) + \dim\left(A^j_{j-1} \cap \Gamma(\Theta_j)\right) \\ &\qquad - \dim\left(A^j_{j-1} \cap \Gamma(\Theta_j) \cap A^j_j\right). \end{aligned}$$

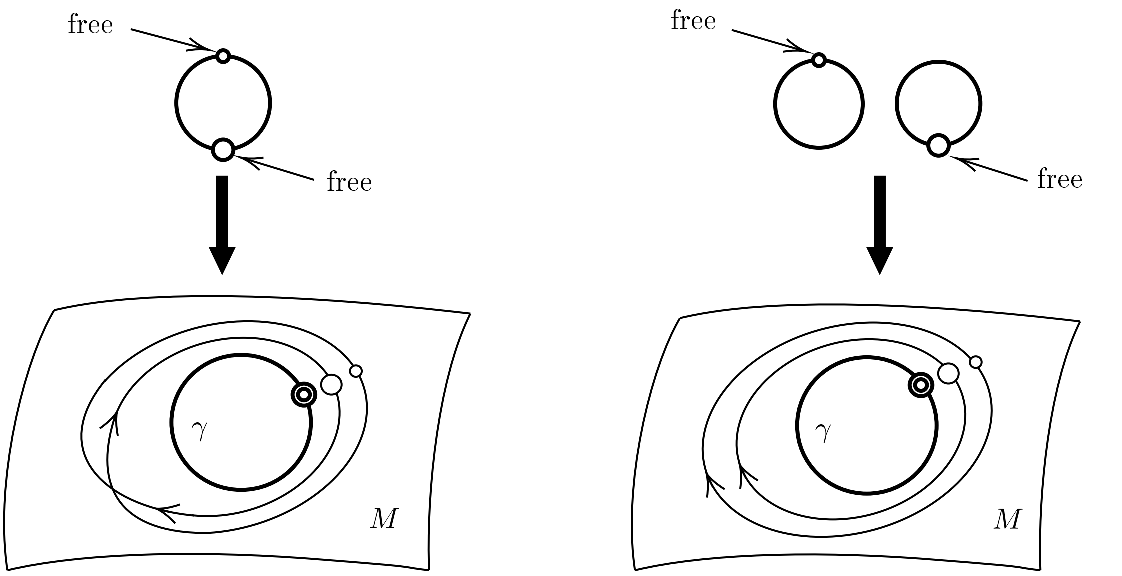
This is the same formula as in [19] modulo terms containing dimensions of intersections.

We end this subsection with a couple of remarks regarding this formula. First of all, in practice the dimensions are reduced even more because $A^j_{j-1} \cap A^j_j \neq \emptyset$. Nevertheless further reductions depend on the structure of the graph and the filtration chosen. Secondly, at first sight it may seem that the formula is a sum of local contributions, because we have used only edges incident to a given vertex in the derivation. However, this is not the case. The non-locality is hidden in the reduced space $V_j^\perp / V_j$ and the corresponding projection π_j.

For example, in the case when $\mathcal{G}_1$ is a tree, we can define a partial order $\leq$ on $\mathcal{G}_0$ by saying that $v \leq w$ if the minimal path from v to the root crosses less or equal number of vertices than the minimal path from w to the root. If we choose a filtration, which orders vertices one by one compatible with the partial ordering, one can identify the set $\mathcal{G}_0^j$ with a sub-tree of $\mathcal{G}$. Then the formula for the indices $i(A^j_{j-1}, \Gamma(\Theta_j), A^j_j)$ will contain terms involving symplectomorphisms of all of the edges in the sub-tree determined by $\mathcal{G}_0^j$ and not only of the incident edges. This can be checked via a long but straightforward calculation.

4.3 Iteration Formulae

We now investigate the case of periodic extremals. We are going to prove two versions of iteration formulae, namely Theorem 4.2 and theorem 4.3 below. Suppose that γ is a closed periodic extremal trajectory. It is straightforward to see that iterations (i.e. concatenation of γ with itself) are still critical points. For the first formula we are going to compute an iterative scheme similar to the one used to prove Theorem 4.1. The following drawing represent the procedure in the case we iterate twice a periodic extremal of period T. We split the interval $[0, 2T]$ into two separate intervals of length $[0, T]$, adding a continuity condition at T gives the variational problem for the original extremal.



We will use the following notation: γ^k will denote the k−th iteration of γ whereas $\operatorname{ind}^- Q_\gamma$ and $\operatorname{ind}^- Q_{\gamma^k}$ the Morse index of γ and γ^k respectively as periodic trajectories. We want to compute the difference $\operatorname{ind}^- Q_{\gamma^k} - k \operatorname{ind}^- Q_\gamma$.

First of all we compute the difference $\operatorname{ind}^- Q_{\gamma^k} - \operatorname{ind}^- Q_{\gamma^{k-1}}$. Let us consider the following manifolds of constraints:

$$\Delta^{\circlearrowleft} := \{(q_1, q_2, q_2, q_1) : q_i \in M, \} \subset M^2 \times M^2,$$
$$\Delta^2 = \{(q_1, q_2, q_1, q_2) : q_i \in M\} \subset M^2 \times M^2.$$

When we restrict to variations satisfying the boundary conditions given by $\Delta^{\circlearrowleft}$, we consider variations of γ^k as a periodic trajectory, whereas when we take boundary conditions Δ^2, we consider independent variations of γ^{k-1} and γ as periodic trajectories, as explained visually in Section 4.3 above. Now we prove the following lemma:

Lemma 4.1. *Here* $n = \dim(M)$, *let* $\Gamma^j = \Gamma(\Theta^j)$ *the graph of* Θ^j. *Then:*

$$\operatorname{ind}^- Q_{\gamma^k} - \operatorname{ind}^- Q_{\gamma^{k-1}} = \operatorname{ind}^- Q_\gamma + i\big(\Gamma^{k-1}, T_{\underline{\lambda}} A(\Delta), \Gamma^k\big) - n + \dim\big(\ker(\Theta^{k-1} - 1)\big).$$

Proof. The statements follows applying Theorem 1.1. We take as $N_1 = \Delta^{\circlearrowleft}$ and as $N_2 = \Delta^2$. The part coming from the dimension is immediate, the intersection of the tangent spaces has dimension n while the dimension of Δ^2 is $2n$. So we get a $-n$.

For the part concerning the intersection between annihilators and graphs, one can check that $T_{(\lambda(0),\lambda(0))}A(\Delta^2) \cap \Gamma(\Theta^{k-1} \times \Theta)$ is isomorphic to the sum of $\ker(\Theta^{k-1} - 1)$ and $\ker(\Theta - 1)$. The triple intersection consists of $\ker(\Theta - 1)$ and thus the term in the statement.

From the definitions it follows that, when we impose the boundary conditions Δ^2, we have $\operatorname{ind}^- Q_{N_2} = \operatorname{ind}^- Q_{\gamma^{k-1}} + \operatorname{ind}^- Q_\gamma$, so the only thing to check is the Maslov index part.

The equation defining the subspace are the following:

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \Theta^{k-1}(\xi_1) \\ \Theta(\xi_2) \end{pmatrix} = \begin{pmatrix} X_1 + Y_1 \\ X_2 + Y_2 \\ X_2 + Y_1 \\ X_1 + Y_2 \end{pmatrix} \qquad X_i, Y_i, \xi_i \in T_{\lambda(0)}(T^*M).$$

By subtracting the second and the third equations, and then the first and the fourth equations we find

$$\xi_2 - \Theta^{k-1}(\xi_1) = Y_2 - Y_1,$$
$$\Theta^{k-1}(\xi_1) - \Theta^k(\xi_2) = \Theta^{k-1}(Y_1 - Y_2).$$

Changing coordinates and setting $\eta = \Theta^{k-1}(\xi_1)$, $Y_1 - Y_2 = \eta_1$ and $\eta_2 = \xi_2$ we get:

$$(\eta, \eta) \in T_{\underline{\lambda}}A(\Delta) \cap (\Gamma^{k-1} + \Gamma^k) \iff \begin{pmatrix} \eta \\ \eta \end{pmatrix} = \begin{pmatrix} \eta_1 + \eta_2 \\ \Theta^{k-1}(\eta_1) + \Theta^k(\eta_2) \end{pmatrix}$$

So we can see that the Maslov form reduces to a form on $\Delta \cap (\Gamma^{k-1} + \Gamma^k)$. Moreover the quadratic form reads:

$$\begin{aligned} m(\xi_1, \xi_2) = \sigma(X_2, Y_1 - Y_2) - \sigma(X_1, Y_1 - Y_2) &= \sigma(\xi_2 - \Theta(\xi_2), Y_1 - Y_2) \\ &= \sigma(\eta_2 - \Theta(\eta_2), \eta_1) \\ &= \sigma(\eta_2, \eta_1) - \sigma(\Theta(\eta_2), \eta_1) \\ &= -\sigma(\eta_1, \eta_2) + \sigma(\Theta^{k-1}(\eta_1), \Theta^k(\eta_2)). \end{aligned}$$

Which is exactly $m(\Gamma^{k-1}, T_{\underline{\lambda}}A(\Delta), \Gamma^k)$ in the coordinates just introduced. And thus the formula is proved. $\square$

The first iteration formula is now a direct consequence of the Lemma just proved:

Theorem 4.2 (Iteration Formulae I). *The index of the k−th iteration of γ as a periodic trajectory satisfies:*

$$\operatorname{ind}^- Q_{\gamma^k} - k \operatorname{ind}^- Q_\gamma = \sum_{j=1}^{k} i\big(\Gamma(\Theta^{j-1}), T_{\underline{\lambda}}A(\Delta), \Gamma(\Theta^j)\big) - \dim(M) + \dim(\ker(\Theta^{j-1} - 1)). \tag{4.3}$$

Proof. We will use an inductive procedure in a similar spirit as in the proof of Theorem 4.1. First we will look at γ^k as the concatenation of γ^{k-1} and γ and express the difference $\operatorname{ind}^- Q_{\gamma^k} - \operatorname{ind}^- Q_\gamma$ in terms of $\operatorname{ind}^- Q_{\gamma^{k-1}}$.

This is the first step of the scheme and was proved in Lemma 4.1. Then we apply the argument iteratively to obtain:

$$\operatorname{ind}^- Q_{\gamma^k} - k \operatorname{ind}^- Q_\gamma = \sum_{j=1}^{k} i(\Gamma^{j-1}, T_{\underline{\lambda}}A(\Delta), \Gamma^j) - n + \dim(\ker(\Theta^{j-1} - 1)).$$

Which is precisely the formula in the statement. □

Now we prove the second iteration formula.

Theorem 4.3 (Iteration Formulae II). *The index of the $k-$th iteration of γ as a periodic trajectory satisfies:*

$$\operatorname{ind}^- Q_{\gamma^k} - k \operatorname{ind}^- Q_\gamma = \sum_{j=1}^{k-1} \dim(M) - \dim(\ker(\Theta - \omega^j)) - i(\Gamma(\Theta), \Delta, \Gamma(\omega^j\Theta)). \quad (4.4)$$

Where ω is a primitive $k-$th root of the unity.

Proof. We work on the space $M^k = M \times \cdots \times M$. The first set of boundary conditions we are going to consider is the following:

$$\Delta^{\circlearrowleft} := \{(q_1, \dots, q_k, r_1 \dots r_k) : r_i, q_i \in M, q_i = r_{i-1}\} \subset M^k \times M^k.$$

Set $q_0 = \gamma(0) = \gamma(1)$. Any curve satisfying the boundary conditions $\Delta^{\circlearrowleft}$ at $(q_0, \dots, q_0)$ gives a variation of the $k-$th iterate of γ seen as periodic trajectory.

The other sets of constraints we are going to introduce are the following:

$$\Delta^k = \{(q_1, \dots, q_k, q_1, \dots, q_k) : q_i \in M\} \subset M^k \times M^k,$$
$$\underline{q}_0 = \{(q_0, \dots, q_0) : q_0 = \gamma(0) = \gamma(1)\}.$$

The first boundary condition is the product of $2k$ copies of the diagonal. Any curve satisfying this set of constraints at point $(q_0, \dots, q_0)$ is a variation of γ^k as k independent periodic trajectories γ. The second boundary condition corresponds to k copies of a single point q_0. Variations of γ^k satisfying these latter conditions are k independent variations of γ as a trajectory with fixed points.

Set for brevity $\Delta^k = T_{\underline{\lambda}}A(\Delta^k)$, $\Delta^{\circlearrowleft} = T_{\underline{\lambda}}A(\Delta^{\circlearrowleft})$ and $\Gamma = \Gamma(\Theta \times \cdots \times \Theta)$ to be the product of k copies of $\Gamma(\Theta)$. We have $T_{\underline{\lambda}}A(\underline{q}_0) = \Pi^{2k}_{\lambda(0)} = \Pi^{2k}$ where $\lambda(0)$ is the initial covector of the lift to the cotangent bundle.

First of all we compute directly $\operatorname{ind}^- Q_{\gamma^k}$ using Theorem 1.1, comparing with the fixed points problem. We get:

$$\operatorname{ind}^- Q_{\gamma^k} = k \operatorname{ind}^- Q_0 + i\big(\Pi^{2k}, \Gamma, \Delta^{\circlearrowleft}\big) + \dim(\Gamma \cap \Pi^{2k}) - \dim(\Gamma \cap \Pi^{2k} \cap A(\Delta^{\circlearrowleft})).$$

Here the notation $\operatorname{ind}^- Q_0$ stands for the index of Q at γ seen as a trajectory with fixed end points. First of all we analyse the term $i\big(\Pi^{2k}, \Gamma, \Delta^{\circlearrowleft}\big)$. To compute it we present the Maslov form as the direct sum of k forms defined on a $n-$dimensional subspace. This is done in Lemma 4.2, where we use the *complexified* version of Maslov index.

The term $i\big(\Pi^{2k}, \Gamma, \Delta^{\circlearrowleft}\big)$ is thus the sum of contributions of the type $i\big(\Pi^2, \Gamma(\omega^j\Theta), \Delta\big)$ where ω is a primitive root of unity.

$$i\big(\Pi^{2k},\Gamma,\Delta^{\circlearrowright}\big)=\sum_{j=0}^{k-1} i\big(\Pi^2,\Gamma(\omega^j\Theta),\Delta\big).$$

Now we apply Theorem 1.1 to the second set of boundary conditions, i.e. Δ^k. We find that:

$$k\,\mathrm{ind}^-\,Q_\gamma=k\,\mathrm{ind}^-\,Q_0+i\big(\Pi^{2k},\Gamma,\Delta^k\big)+\dim(\Gamma\cap\Pi^{2k})-\dim(\Gamma\cap\Pi^{2k}\cap\Delta^k).$$

Exactly as in the previous case the piece $i\big(\Pi^{2k},\Gamma,\Delta^k\big)$ splits as a sum. But this time the reason is more apparent: we are considering independent variation on each iteration. It follows that $i\big(\Pi^{2k},\Gamma,\Delta^k\big)=k\,i\big(\Pi^2,\Gamma(\Theta),\Delta\big)$

Now we subtract the two equations and we are left with the following expression for $\mathrm{ind}^-\,Q_{\gamma^k}-k\ \mathrm{ind}^-\,Q_\gamma$:

$$\begin{aligned}\mathrm{ind}^-\,Q_{\gamma^k}-k\ \mathrm{ind}^-\,Q_\gamma=\sum_{j=0}^{k-1}&\Big(i\big(\Pi^2,\Gamma(\omega^j\Theta),\Delta^{\circlearrowright}\big)-i\big(\Pi^2,\Gamma(\Theta),\Delta\big)\Big)+\\&+\dim(\Gamma\cap\Pi^{2k}\cap\Delta^k)-\dim(\Gamma\cap\Pi^{2k}\cap\Delta^{\circlearrowright}).\end{aligned}$$

Let's rewrite the term involving the intersections. It is straightforward to see that $\dim(\Gamma\cap\Pi^{2k}\cap\Delta^k)=k\dim(\Gamma(\Theta)\cap\Pi^2\cap\Delta)$. In turn this can be easily seen to be $k\dim(\ker(\Theta-1)\cap\Pi)$.

For the second piece it holds that:

$$\dim(\Gamma\cap\Pi^k\cap\Delta^{\circlearrowright})=\sum_{j=0}^{k-1}\dim(\ker(\Theta-\omega^j)\cap\Pi).$$

We prove this below, in Proposition 4.2. Putting all together we get:

$$\dim(\Gamma\cap\Pi^{2k}\cap\Delta^k)-\dim(\Gamma\cap\Pi^k\cap\Delta^{\circlearrowright})=\sum_{j=0}^{k-1}\dim(\Gamma(\Theta)\cap\Pi^2\cap\Delta)-\dim(\ker(\Theta-\omega^j)\cap\Pi).\quad(4.5)$$

Now we can use the cocycle property given in eq. (A.7) with the subspaces Π, $\Gamma(\omega^j\Theta)$, $\Gamma(\Theta)$ and Δ to express the terms in the sum using the Maslov index of the spaces $\Gamma(\omega^j\Theta)$ and Δ. These computations are collected in Proposition 4.2. What we find is that:

$$\begin{aligned}i\big(\Pi^2,\Gamma(\omega^j\Theta),\Delta\big)-i\big(\Pi^2,\Gamma(\Theta),\Delta\big)=-i\big(\Gamma(\Theta),\Delta,\Gamma(\omega^j\Theta)\big)-\dim(\ker(\Theta-1)\cap\Pi)+\\+\dim(\ker(\Theta-\omega^{-j})\cap\Pi)-\dim\ker(\Theta-\omega^{-j})+\dim(M).\end{aligned}$$

Since we are summing over $j=0,\dots,k-1$ and ω is a primitive k-th root of unity, we have that $\sum_{j=0}^{k-1}\dim(\ker(\Theta-\omega^{-j})\cap\Pi)=\sum_{j=0}^{k-1}\dim(\ker(\Theta-\omega^{j})\cap\Pi)$ and thus the intersection of the eigenspaces with the fibre cancel out with the part coming from triple intersection given in eq. (4.5).

Summing up we finally obtain:

$$\mathrm{ind}^-\,Q_{\gamma^k}-k\ \mathrm{ind}^-\,Q_\gamma=\sum_{j=1}^{k-1}\dim(M)-\dim\ker(\Theta-\omega^j)-i\big(\Gamma(\Theta),\Delta,\Gamma(\omega^j\Theta)\big).$$

Which is exactly the statement of the theorem. □

Lemma 4.2. *Let $\omega \in \mathbb{C}$ be a primitive $k-$th root of* 1*. The Maslov form $m(\Pi^{2k}, \Gamma, \Delta^{\circlearrowright})$ splits a direct sum $\bigoplus_{i=0}^{k-1} m_i$ where:*

$$m_i = m(\Pi^2, \Gamma(\omega^i \Theta), \Delta).$$

Proof. We will use the Hermitian version of Maslov index. Any real subspace V appearing in the proof will stand for its complexification $V \otimes \mathbb{C}$ without any mention of the tensor product operation. Let us write down the equation defining the space $(\Pi^{2k} + \Delta^{\circlearrowright}) \cap \Gamma$.

$$v \in \Gamma \iff v = (\xi_1, \dots, \xi_k, \Theta(\xi_1), \dots, \Theta(\xi_k)), \quad \xi_j \in T_{\lambda_0}(T^*_{q_0}M).$$

On the other hand belonging to $\Pi^{2k} + \Delta^{\circlearrowright}$ means:

$$v \in \Pi^{2k} + \Delta^{\circlearrowright} \iff v = (\mu_1, \dots, \mu_k, \nu_1, \dots, \nu_k), \quad \mu_{i+1} - \nu_i \in \Pi,$$

where $\mu_{k+1} = \mu_1$. So the space $(\Pi^{2k} + \Delta^{\circlearrowright}) \cap \Gamma$ is given by $\{(\xi_1, \dots, \xi_k) : \xi_{i+1} - \Theta(\xi_i) \in \Pi\}$.

Maslov form is computed in the following way. Let

$$\begin{aligned} \xi_i &= X_i + \alpha_i, \\ \Theta(\xi_i) &= X_{i+1} + \beta_i, \end{aligned}$$

where $\alpha_i, \beta_i \in \Pi$, $X_i \in T_{\lambda_0}(T^*M)$. Then we have

$$\begin{aligned} m(\underline{\xi}) = \sum_{i=1}^{k} -\sigma(\bar{\alpha}_i, X_i) + \sigma(\bar{\beta}_i, X_{i+1}) &= \sum_{i=1}^{k} -\sigma(\bar{\alpha}_i, X_i) + \sigma(\bar{\beta}_{i-1}, X_i) \\ = \sum_{i=1}^{k} \sigma(-\bar{\alpha}_i + \bar{\beta}_{i-1}, X_i) &= \sum_{i=1}^{k} \sigma(-\bar{\alpha}_i + \bar{\beta}_{i-1}, \xi_i) \\ = \sum_{i=1}^{k} \sigma(\Theta(\bar{\xi}_{i-1}) - \bar{\xi}_i, \xi_i) &= \sum_{i=1}^{k} \sigma(\Theta(\bar{\xi}_{i-1}), \xi_i) - \sigma(\bar{\xi}_i, \xi_i). \end{aligned}$$

Where in the third equality we simply shifted the second index cyclically.

Suppose that ω is a primitive $k-$th root of the identity and make the following change of variables.

$$\underline{\xi} = (\xi_1, \dots, \xi_k) \mapsto \left(\sum_{i=1}^{k} \xi_i, \dots, \sum_{i=1}^{k} \omega^{j(i-1)} \xi_i, \dots, \sum_{i=1}^{k} \omega^{(k-1)(i-1)} \xi_i \right) =: \underline{\eta},$$

which essentially is just the Kronecker product of the identity with the transpose of Vandermonde's matrix obtained with $\{1, \omega, \dots, \omega^{k-1}\}$. In the new coordinates the equation reads:

$$\begin{aligned} \eta_l - \omega^{l-1}\Theta(\eta_l) &= \sum_{i=1}^{k} \omega^{(l-1)(i-1)} \xi_i - \sum_{i=1}^{k} \omega^{(l-1)i} \Theta(\xi_i) \\ &= \sum_{i=1}^{k} \omega^{(l-1)i} \xi_{i+1} - \omega^{(l-1)i} \Theta(\xi_i) \\ &= \sum_{i=1}^{k} \omega^{(l-1)i} (\xi_{i+1} - \Theta(\xi_i)) \in \Pi. \end{aligned}$$

So in the new coordinates the space $(\Pi^{2k}+\Delta^{\circlearrowleft})\cap\Gamma$ splits as the direct sum $\bigoplus_{l=1}^{k}\{\eta : \eta-\omega^l\Theta(\eta)\in\Pi\}$.

The inverse transformation is given by the following rule:

$$\xi_i=\frac{1}{k}\sum_{l=1}^{k}\omega^{-(i-1)(l-1)}\eta_l.$$

If we plug in the second term of the Maslov form we have:

$$\sum_{i=1}^{k}\sigma(\Theta(\overline{\xi_i}),\xi_{i+1})=\frac{1}{k^2}\sum_{i,l,s=1}^{k}\sigma(\Theta(\omega^{(i-1)(s-1)}\bar{\eta}_s),\omega^{-i(l-1)}\eta_l)$$

$$=\frac{1}{k^2}\sum_{i,l,s=1}^{k}\omega^{i(s-l)}\omega^{-(s-1)}\sigma(\Theta(\bar{\eta}_s),\eta_l)=\frac{1}{k^2}\sum_{l,s=1}^{k}\Big(\sum_{i=1}^{k}(\omega^{i(s-l)})\omega^{1-s}\sigma(\Theta(\bar{\eta}_s),\eta_l)\Big).$$

In particular the only non zero terms are those for which $s=l$ since the sum of powers of any primitive root (up to k) is zero.

We can handle similarly the first term. In this way we find that the Maslov form on our subspace splits in the following way:

$$m(\underline{\eta})=\frac{1}{k}\sum_{s=1}^{k}\sigma(\overline{\omega^{s-1}\Theta(\eta_s)},\eta_s)-\sigma(\bar{\eta}_s,\eta_s).$$

The factor $\frac{1}{k}$ is irrelevant for us and comes just from the change of coordinates. The last step is to identify the summands with $m(\Pi^2,\Gamma(\omega^{s-1}\Theta),\Delta)$. Let's write down the kernel for these forms. The space we have to look at is $(\Pi^2+\Delta)\cap\Gamma(\omega^{s-1}\Theta)$. It is defined by:

$$\eta=\alpha+X\quad\omega^{s-1}\Theta(\eta)=\beta+X\quad\alpha,\beta\in\Pi.$$

By the definition the Maslov form is given by

$$m(\eta)=-\sigma(\bar{\alpha},X)+\sigma(\bar{\beta},X)=\sigma(\overline{\omega^{s-1}\Theta(\eta)}-\bar{\eta},\eta).$$

$\square$

Proposition 4.2. *The following relation holds:*

$$i\big(\Pi^2,\Gamma(\omega^j\Theta),\Delta\big)-i\big(\Pi^2,\Gamma(\Theta),\Delta\big)=-i\big(\Gamma(\Theta),\Delta,\Gamma(\omega^j\Theta)\big)+\dim(M)+d_j,\tag{4.6}$$

where $d_j=-\dim(\ker(\Theta-1)\cap\Pi)+\dim(\ker(\Theta-\omega^{-j})\cap\Pi)-\dim\ker(\Theta-\omega^{-j})$. *Moreover the space* $\Gamma\cap\Pi^{2k}\cap\Delta^{\circlearrowleft}$ *splits as a direct sum and its dimension is given by:*

$$\dim(\Gamma\cap\Pi^{2k}\cap\Delta^{\circlearrowleft})=\sum_{j=0}^{k-1}\dim(\ker(\Theta-\omega^j)\cap\Pi).$$

Proof. The second part can be deduced by the proof of Lemma 4.2. In fact the space $\Pi^{2k}\cap\Gamma\cap\Delta^{\circlearrowleft}$ is isomorphic to $\bigoplus_i\ker(\Theta-\omega^i)\cap\Pi$. This can be either directly computed from the definition of the spaces or deduced in the following way.

Let P represent the standard $k-$cycle which maps $\xi_i \to \xi_{i+1}$ and $\xi_k \to \xi_1$. A direct calculation shows that $\Delta^{\circlearrowright} = \Gamma(P)$. Thus any element of $\Delta^{\circlearrowright} \cap \Gamma$ can be written as

$$\begin{pmatrix} \xi \\ P\xi \end{pmatrix} = \begin{pmatrix} \xi \\ diag(\Theta)(\xi) \end{pmatrix} \iff P^{-1}diag(\Theta)(\xi) = \xi \iff diag(\Theta)P^{-1}(\eta) = \eta,\, \eta = P\xi.$$

i.e. an eigenvalue problem.

The core of the proof of Lemma 4.2 consisted in the diagonalization of the following matrix:

$$\begin{pmatrix} \Theta & & \\ & \ddots & \\ & & \Theta \end{pmatrix} P^{-1} \sim \begin{pmatrix} \omega^0\Theta & & \\ & \ddots & \\ & & \omega^{k-1}\Theta \end{pmatrix}$$

with the remaining elements understood to be zero. The transformation diagonalizing the matrix we used preserves the fibre. So it follows that $\Pi^{2k} \cap \Gamma \cap \Delta^{\circlearrowright}$ is the sum of the eigenspaces $\ker(\Theta - \omega^j)$ intersected with the fibre Π.

Now we prove the first part of the proposition. Let us apply the cocycle property to the subspaces Π^2, $\Gamma(\omega^j\Theta)$, $\Gamma(\Theta)$ and Δ.

$$\begin{aligned} i\big(\Pi^2, \Gamma(\omega^j\Theta), \Delta\big) - i\big(\Pi^2, \Gamma(\Theta), \Delta\big) &= i\big(\Gamma(\Theta), \Pi^2, \Gamma(\omega^j\Theta)\big) \\ &\quad - i\big(\Gamma(\Theta), \Delta, \Gamma(\omega^j\Theta)\big) + c_i, \\ c_i = \dim(\Theta(\Pi) \cap \Pi) - \dim(\ker(\Theta - 1) &\cap \Pi) + \dim(\ker(\Theta - \omega^{-j}) \cap \Pi) + \\ - \dim\ker(\Theta - \omega^{-j}). \end{aligned}$$

The formula is almost the one given in the statement except for the terms $\dim(\Theta(\Pi)\cap\Pi)$ and $i\big(\Gamma(\Theta), \Pi^2, \Gamma(\omega^j\Theta)\big)$ and a lacking $\dim(M)$.

We can compute the Maslov index term in the following way. Notice that $\Gamma(\omega^j\Theta)$ and $\Gamma(\omega^l\Theta)$ are transversal if the index j is different form l. It follows that the space on which the form is defined is Π^2. Moreover the equations are $\xi_1 + \xi_2 = \nu_1 \in \Pi$ and $\Theta(\xi_1 + \omega^j\xi_2) = \nu_2 \in \Pi$. Thus Maslov form reads:

$$m(\nu_1, \nu_2) = -\sigma(\bar{\xi}_1, \xi_2) + \omega^j\sigma(\Theta(\bar{\xi}_1), \Theta(\xi_2)) = (\omega^j - 1)\sigma(\bar{\xi}_1, \xi_2).$$

We can invert the equations to write them on Π^2. We get $\xi_2 = \frac{1}{1-\omega^j}(\nu_1 - \Theta^{-1}(\nu_2))$ and $\xi_1 = \frac{1}{1-\omega^j}(\Theta^{-1}(\nu_2) - \omega^j\nu_1)$ and thus the form is equivalent to:

$$m(\nu_1, \nu_2) = \frac{1}{\bar{\omega}^j - 1}(\sigma(\bar{\nu}_2, \Theta(\nu_1)) + \omega^{-j}\sigma(\bar{\nu}_1, \Theta^{-1}(\nu_2))).$$

This form has zero signature and kernel isomorphic to two copies of $\Theta(\Pi) \cap \Pi$. This is a general fact and can be seen as follow. Suppose the matrix representing the quadratic form has the following expression:

$$\mathcal{M} = \begin{pmatrix} 0 & X \\ \bar{X}^* & 0 \end{pmatrix} \qquad m(\nu_1, \nu_2) = \langle \bar{\nu}_1, X\nu_2 \rangle + \langle \bar{\nu}_2, \bar{X}^*\nu_1 \rangle.$$

Let be Q and R unitary matrices which gives the singular values decomposition for X, i.e. $QXR = D$ for $D = diag(d_i^2)$, diagonal and with non negative entries.

Apply the following change of coordinates to $\mathcal{M}$:

$$\begin{pmatrix} Q & 0 \\ 0 & \bar{R}^* \end{pmatrix} \begin{pmatrix} 0 & X \\ \bar{X}^* & 0 \end{pmatrix} \begin{pmatrix} \bar{Q}^* & 0 \\ 0 & R \end{pmatrix} = \begin{pmatrix} 0 & D \\ D & 0 \end{pmatrix}.$$

And then apply another change:

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & D \\ D & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -2D & 0 \\ 0 & 2D \end{pmatrix}.$$

Thus the non zero eigenvalues of the matrix $\mathcal{M}$ are $\pm d_i^2$, where $d_i^2 > 0$ are the positive singular values of X. The kernel of $\mathcal{M}$ has dimension $2\dim\ker(X)$.

This is precisely our situation: fix a Lagrangian complement to the fibre Π, and consider the matrices associated to Θ and $\frac{\omega^{-j}}{\omega^{-j}-1} J\Theta^{-1}$. In blocks they can be written as:

$$\Theta = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad J\Theta^{-1} = \begin{pmatrix} C^* & -A^* \\ D^* & -B^* \end{pmatrix}, \quad J\Theta = \begin{pmatrix} -C & -D \\ A & B \end{pmatrix}.$$

We are using coordinates in which the fibre Π is the span of the first n coordinates. Thus the block we have to consider is always the upper left one. Our form, with this conventions, is written as:

$$m(\nu_1, \nu_2) = \left\langle \bar{\nu}_2, \frac{1}{1-\omega^{-j}} C\nu_1 \right\rangle + \left\langle \bar{\nu}_1, \frac{\omega^{-j}}{\omega^{-j}-1} C^*\nu_2 \right\rangle.$$

So for us $X = \frac{\omega^{-j}}{1-\omega^{-j}} C^*$. Thus our form has zero signature, is defined on a $2\dim(M)$ dimensional vector space and the kernel is isomorphic to two copies of the kernel of X. The latter is easily seen to be $\Theta(\Pi) \cap \Pi$.

Thus it follows that $i\big(\Gamma(\Theta), \Pi^2, \Gamma(\omega^j\Theta)\big) = \dim(M) - \dim(\Theta(\Pi) \cap \Pi)$. Inserting above we get the formula in the statement. □

We can consider the function $\mathbb{S}^1 \ni z \mapsto i(\Gamma(\Theta), \Delta, \Gamma(z\Theta))$. It has very nice properties and an explicit description in terms of the monodromy matrix Θ. These ideas are collected in the following proposition.

Proposition 4.3. *The number $i\left(\Gamma(\Theta), \Delta, \Gamma(\omega^j\Theta)\right)$ corresponds to the number of negative eigenvalues of the following matrix:*

$$M_{\omega^j} = \frac{1}{1-\omega^{-j}} J\Big(\omega^{-j} + 1 - \omega^{-j}\Theta - \Theta^{-1}\Big).$$

If we consider the function $\mathbb{S}^1 \ni z \mapsto i\big(\Gamma(\Theta), \Delta, \Gamma(z\Theta)\big)$, it is locally constant with at most $2n$ jumps at eigenvalues of Θ. Moreover the jumps are bounded in amplitude by $\dim(\ker(\Theta - z))$ where $z \in \mathbb{S}^1$.

Proof. The first part is just a straightforward computation. Take for any $\alpha \in \mathbb{S}^1$:

$$\begin{pmatrix} \xi_1 \\ \Theta(\xi_1) \end{pmatrix} + \begin{pmatrix} \xi_2 \\ \alpha\Theta(\xi_2) \end{pmatrix} = \begin{pmatrix} X \\ X \end{pmatrix} \Rightarrow \begin{cases} (1-\alpha)\xi_2 = X - \Theta^{-1}(X), \\ (\alpha - 1)\xi_1 = \alpha X - \Theta^{-1}(X). \end{cases}$$

If $\alpha \neq 1$ the two graphs are always transversal and the Maslov quadratic form can be written in terms of the variable X:

$$\begin{aligned} m(X) &= -\sigma(\bar{\xi}_1, \xi_2) + \alpha\sigma(\Theta(\bar{\xi}_1), \Theta(\xi_2)) = (\alpha - 1)\sigma(\bar{\xi}_1, \xi_2) \\ &= \frac{1}{1-\bar{\alpha}}\sigma\Big(\bar{\alpha}\bar{X} - \Theta^{-1}(\bar{X}), X - \Theta^{-1}(X)\Big) \\ &= \frac{\bar{\alpha}+1}{1-\bar{\alpha}}\sigma(\bar{X}, X) - \frac{1}{1-\bar{\alpha}}\sigma((\bar{\alpha}\Theta + \Theta^{-1})(\bar{X}), X). \end{aligned}$$

It follows that the kernel is $M_\alpha = \frac{1}{1-\bar{\alpha}} J\Big(\bar{\alpha} + 1 - \bar{\alpha}\Theta - \Theta^{-1}\Big)$.

For the second part notice that the map $\alpha \mapsto M_\alpha$ is continuous away from 1 with values in the space of Hermitian matrices.

A change of index can occur only at those points in which the determinant of M_z is zero, thus at most $2n$ times. Moreover the jumps are the following:

$$\det(M_\alpha) = 0 \iff \det(\bar{\alpha} + 1 - \bar{\alpha}\Theta - \Theta^{-1}) = 0,\ \alpha \neq 1.$$

In particular, notice that Θ and Θ^{-1} can be put in the same block triangular form. For example one can choose to put Θ in its Jordan form. On the diagonal, at a block corresponding to eigenvalue λ of Θ, the elements are $\bar{\alpha} + 1 - \bar{\alpha}\lambda - \frac{1}{\lambda}$. This quantity is zero if and only if $\alpha = \frac{\lambda}{|\lambda|^2}$ i.e. if α is an eigenvalue of Θ that lies on the circle.

Thus the jumps are at most $2n$. The part on the bound follows by this observation: take a Jordan block of Θ with eigenvalue λ. Then the corresponding block of Θ^{-1} will have $\bar{\lambda}$ on the diagonal and $(-1)^k\bar{\lambda}^{k+1}$ on the $k-$th upper diagonal. This implies that on the first upper diagonal of the λ block of $\bar{\lambda}\Theta + \Theta^{-1}$ we considered you end up with $-\bar{\lambda} + \bar{\lambda}^2$, which is different from zero. Thus each $\lambda-$block contributes with a single eigenvalue and so the jumps are controlled by $\dim(\ker(\Theta - \lambda))$. □

4.4 Second Variation with moving endpoints

Let $\lambda : [0,1] \to T^*M$ be an extremal satisfying PMP for the problem in eq. (B.2) with $N = N_0 \times N_1$. Denote by $\tilde{u}$ the corresponding control and by $\gamma(t) = \pi(\lambda(t))$, $t \in [0,1]$ the extremal curve on the manifold M. As discussed in Section 2.3 we can extend γ to an admissible curve of the auxiliary system in eq. (2.5)

$$\hat{\gamma} = \begin{cases} q_0, & \text{if } t < 0, \\ \gamma(t), & \text{if } t \in [0,1], \\ q_1, & \text{if } t > 1, \end{cases} \qquad \hat{\tilde{u}}(t) = \begin{cases} 0, & \text{if } t < 0, \\ \tilde{u}(t), & \text{if } t \in [0,1], \\ 0, & \text{if } t > 1. \end{cases}$$

In order to simplify slightly the notations, we will omit in the future the hat symbol for $\hat{\tilde{u}}$ by essentially identifying $\tilde{u} \sim (0, \tilde{u}, 0)$.

In Section 2.3 we computed the first and second variations at a critical point $\tilde{u}$. To do this, we used the already existing formulas for the fixed end-point problem which can be found in several references such as [12] but applied to the auxiliary system in eq. (2.5).

Recall that $\Pi := \Pi_{\lambda_0}$ denotes the vertical subspace, namely the tangent space to fibre T_q^*M described in eq. (A.9). The kernel of the differential of the endpoint mapping and the second variation are:

$$\ker d_{\tilde{u}}E = \left\{ \hat{v} \in \mathbb{R}^{\dim N_0} \oplus L^\infty([0,1], \mathbb{R}^k) \oplus \mathbb{R}^{\dim N_1} : \int_{-1}^{2} \hat{Z}_t \hat{v}(t) dt \in \Pi \right\}, \tag{4.7}$$

$$Q(\hat{v},\hat{w}) = \int_{-1}^{2} \big[-H_t(\hat{v}(t),\hat{w}(t)) - \int_{-1}^{t} \sigma(\hat{Z}_\tau \hat{v}(\tau), \hat{Z}_t \hat{w}(t)) d\tau \big] dt, \tag{4.8}$$

where $\hat{v},\hat{w} \in \ker d_{\tilde{u}}E$. We can expand the expressions for the first and second variations knowing the particular form of $\hat{Z}_t$. We split the integrals into three integrals over the intervals $[-1,0]$, $[0,1]$ and $[1,2]$ and simplify the integrands using the skew-symmetry of σ (see Section 2.3 and definition 2.1). We obtained the equalities:

$$\ker d_{\tilde{u}}E = \left\{ v \in L^\infty[0,1], v_i \in \mathbb{R}^{\dim N_i} : \int_0^1 Z_t v(t) dt + Z_0 v_0 + Z_1 v_1 \in \Pi \right\}, \tag{4.9}$$

$$\begin{aligned} Q(\hat{v},\hat{w}) = \int_0^1 \Big[-H_t(v(t),w(t)) - \sigma\left(Z_0 v_0 + \int_0^t Z_\tau v(\tau), Z_t w(t) \right) d\tau \Big] dt \\ -\sigma\left(Z_0 v_0 + \int_0^1 Z_t v(t) dt, Z_1 w_1 \right). \end{aligned} \tag{4.10}$$

We are going to work now with this quadratic form and prove Theorem 1.1.

4.4.1 Jacobi equation and Second Variation

For brevity denote $\mathcal{V} = \ker d_{\tilde{u}}E$. Inside $\mathcal{V}$ we look at a distinguished subspace

$$V = \{\hat{v} \in \mathcal{V} : v_0 = 0, v_1 = 0\}, \tag{4.11}$$

which corresponds to variations that fix the end-points q_0, q_1 of an extremal curve γ at first order. More precisely, they constitute the tangent space to the manifold of controls fixing the end-points of γ. Hence $Q|_V$ is the second variation of the optimal control problem with fixed end-points and there exist efficient ways of computing the index of this quadratic forms using generalisations of classical Jacobi fields [12, Section 21]. Our goal is to compute the difference

$$\operatorname{ind} Q - \operatorname{ind} Q|_V$$

in terms of geometric objects on the manifold M, which will result in formula eq. (1.3) when $N = N_0 \times N_1$. The main tool for computing the difference of indices is the following folklore lemma.

Lemma 4.3. *Suppose that Q is a continuous quadratic form on a Hilbert space. Then for any subspace V of finite codimension it holds:*

$$\operatorname{ind} Q = \operatorname{ind} Q|_V + \operatorname{ind} Q|_{V^{\perp_Q}} + \dim \left(V \cap V^{\perp_Q} / (V \cap \ker Q) \right). \tag{4.12}$$

For us the Hilbert space will be the subspace $\mathcal{V}$. Thus

$$V^{\perp_Q} = \{\hat{v} \in \mathcal{V} : Q(\hat{v},\hat{w}) = 0, \forall \hat{w} \in V\},$$

and the kernel of Q on $\mathcal{V}$ is

$$\ker Q = \{\hat{v} \in \mathcal{V} : Q(\hat{v},\hat{w}) = 0, \forall \hat{w} \in \mathcal{V}\}.$$

In the following discussion we derive boundary value problems whose differential equation is a *Jacobi equation*. The solution to those boundary value problems encode all the information about each term in formula (4.12).

By definition the subspace V is the set of variations $\hat{v} = (v_0, v, v_1)$ such that:

$$\int_0^1 Z_t v(t)dt \in \Pi, \text{ and } v_0 = 0,\ v_1 = 0$$

. Moreover, since Π is a Lagrangian subspace

$$\int_0^1 Z_t v(t)dt \in \Pi \iff \sigma\left(\int_0^1 Z_t v(t)dt, \nu\right) = 0,\ \forall\, \nu \in \Pi$$

and so:

$$\begin{aligned} \hat{v} \in V^{\perp_Q} &\iff Q(\hat{v}, \hat{w}) = 0,\ \forall\, \hat{w} \in V \\ &\iff Q(\hat{v}, \hat{w}) = \int_{-1}^2 \sigma(\nu, \hat{Z}_t \hat{w}(t))dt, \quad \forall \nu \in \Pi, \forall\, \hat{w} \in V. \end{aligned}$$

Using formula (4.10) we have that for almost every $t \in [0,1]$ and a vector $\nu \in \Pi$:

$$H_t(v(t), \cdot) + \sigma\left(\int_0^t Z_\tau v(\tau)d\tau + Z_0 v_0, Z_t \cdot\right) = \sigma(Z_t \cdot, \nu).$$

By the strong Legendre condition H_t is invertible. This allows us solve the equation for the variation v and obtain

$$v(t) = H_t^{-1}\sigma\left(Z_t \cdot, \int_0^t Z_\tau v(\tau)d\tau + Z_0 v_0 + \nu\right).$$

Set

$$\eta(t) = \int_0^t Z_\tau v(\tau)d\tau + Z_0 v_0 + \nu.$$

Differentiating η and plugging in the expression for the variation v shows that η satisfies the following equation for almost all $t \in [0,1]$:

$$\dot{\eta}(t) = Z_t H_t^{-1}\sigma(Z_t \cdot, \eta(t)). \tag{4.13}$$

This equation is known as the *Jacobi equation* [12, Theorem 21.1]. Using the definition of Z_0 and Lemma lemma 2.2 we find that $\eta(t)$ satisfies (4.13) with $\pi_*\eta(0) \in T_{q_0}N_0$. To obtain conditions at $t = 1$, we have to use the fact that $\hat{v} \in \mathcal{V}$. In this case from (4.9) it follows that there exists $\xi \in \Pi$ such that

$$\eta(1) = \int_0^1 Z_\tau v(\tau)d\tau + Z_0 v_0 + \nu = \xi + \nu - Z_1 v_1.$$

Thus the variation $\hat{v} \in V^{\perp_Q}$ defines a function $\eta : [0,1] \to T_{\lambda(0)}(T^*M)$ which solves the following boundary value problem

$$\begin{cases} \dot{\eta}(t) = Z_t H_t^{-1}\sigma(Z_t \cdot, \eta(t)), \\ \pi_*\eta(0) \in T_{q_0}N_0, \quad \pi_*\eta(1) \in (\pi \circ \tilde{\Phi}^{-1})_*(T_{q_1}A(N_1)). \end{cases} \tag{4.14}$$

The second space in the boundary condition is the pull back of N_1 to q_0 using the flow of the control system with the extremal control $\tilde{u}$. From this it is immediate to compute the dimension of $V \cap V^{\perp_Q}$. It is enough to substitute $v_i = 0$ in the above equations and thus consider solution starting from Π and arriving to Π. Since the Jacobi equation derived above is exactly the same as the Jacobi equation for problem with fixed points, we immediately see that $\dim(V \cap V^{\perp_Q})$ is the multiplicity of the point q_1 as conjugate point.

In the same spirit we can compute the dimension of $\ker Q \cap V$ using the Jacobi equation. We have

$$\ker Q \cap V = \{\hat{v} \in V : Q(\hat{v}, \hat{w}) = 0, \ \forall \hat{w} \in \mathcal{V}\}.$$

Using the same argument as above we find that for every $\nu \in \Pi$

$$\begin{cases} 0 = \sigma(Z_0 \cdot, \nu) \\ Q(v, \cdot) = \sigma(Z_t \cdot, \nu) \\ \sigma\left(\int_0^1 Z_t v(t) dt, Z_1 \cdot\right) = \sigma(Z_1 \cdot, \nu) \end{cases}$$

The second equation allows us to recover a solution of the Jacobi equation η using the same argument as above, when we considered variations $\hat{v} \in V \cap V^{\perp_Q}$. The first equality gives us a condition on ν and consequently on $\eta(0)$, while the third condition give us a condition for $\eta(1)$. Namely

$$\eta(0) \in \Pi \cap T_{\lambda(0)} A(N_0), \qquad \eta(1) \in \Pi \cap \tilde{\Phi}_*^{-1} T_{\lambda(1)} A(N_1).$$

The following proposition collects all the facts proved above and clarifies the correspondence between controls and solutions of the boundary value problems.

Proposition 4.4. *Consider system (4.13), to any solution η satisfying the boundary value problem (4.14) we can associate a variation $v \in V_Q^{\perp}$ such that $\dot{\eta}(t) = Z_t v(t)$ and vice-versa, modulo solutions satisfying $\eta(0), \eta(1) \in \Pi$ and $\dot{\eta} = 0$. Moreover:*

i) elements inside $V \cap V^{\perp_Q}$ correspond to solutions of (4.13) satisfying the boundary conditions:

$$\eta(0) \in \Pi, \quad \eta(1) \in \Pi;$$

ii) elements of $\ker Q \cap V$ correspond to solutions satisfying the boundary conditions:

$$\eta(0) \in \Pi \cap T_{\lambda(0)} A(N_0), \quad \eta(1) \in \Pi \cap \tilde{\Phi}_*^{-1} T_{\lambda(1)} A(N_1);$$

iii) elements in $\ker Q$ correspond to solutions of (4.13) satisfying the boundary conditions:

$$\eta(0) \in T_{\lambda(0)} A(N_0), \quad \eta(1) \in \tilde{\Phi}_*^{-1} T_{\lambda(1)} A(N_1).$$

Proof. Following the derivation of the Jacobi equation and associated boundary conditions it only remains to prove the first part by computing the kernel of the map $\eta \mapsto v$. If $\hat{v} = 0$, then $v \equiv 0$ for all $t \in [0,1]$ and $v_0 = 0$, $v_1 = 0$. This implies that $\eta \equiv \nu$.

Vice versa, let η be constant with $\eta(0) \in \Pi$. Since Z_i are injective and have non-trivial projections to $T_{\pi(\lambda(0))} M$, it follows that $v_i = 0$. Moreover $\dot{\eta}(t) = Z_t v(t) = 0$ and consequently, by definition of $V^{\perp_Q}$,

$$0 = Q(\hat{v}, \hat{w}) = \int_0^1 H_t(v(t), w(t)) dt, \qquad \forall \hat{w} \in V.$$

In particular $H_t(v(t), v(t)) = 0$ for almost every $t \in [0,1]$. But then by the strong Legendre condition $v \equiv 0$. Point (iii) can be obtained in a similar fashion as point (i) and (ii). $\square$

As before we will denote by Θ the symplectomorphism $\tilde{\Phi}_* \circ \Phi$, where $\tilde{\Phi}_*$ is the differential of the re-parametrization flow given in eq. (B.11) and Φ the fundamental solution of eq. (4.13). As usual $\Gamma(\Theta)$ will stand for the graph of Θ.

Remark 4.3. It can be shown that (4.13) is closely related to the linearisation of the extremal flow along the fixed extremal λ we are considering (when the latter is defined), see for example [13]. It is the linearisation at $\lambda(0)$ of the Hamiltonian flow of $b^t_{\tilde{u}}(\lambda) = (H - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_t(\lambda)$ which coincides with the linearisation of $(\tilde{\Phi}_t)^{-1} \circ e^{tH}$. Let us denote by Φ_t the flow of the Jacobi equation (4.13) at time t and let

$$\Gamma(\Phi_t) = \{(\eta(0), \eta(t)) : \eta(0) \in T_{\lambda(0)}(T^*M)\} \subset T_{\lambda(0)}(T^*M) \times T_{\lambda(0)}(T^*M)$$

be its graph. Then in this notation

$$\Gamma(\Theta_t) = (I \times \tilde{\Phi}_t)_* \Gamma(\Phi_t).$$

We can now compute the restriction of Q to $V^{\perp_Q}$ and prove the following result.

Proposition 4.5. *Let Q be the quadratic form of second variation for the problem in eq. (B.2) and V be the subspace of variations (4.11). Then*

$$\operatorname{ind}^- Q = \operatorname{ind}^- Q|_V + i\big(\Pi^2_{\underline{\lambda}}, \Gamma(\Theta), T_{\underline{\lambda}}A(N)\big) + \dim(\Gamma(\Theta) \cap \Pi^2_{\underline{\lambda}}) - \dim(\Gamma(\Theta) \cap \Pi^2_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(N))$$

Moreover, the Maslov index of the triple can be replaced by $i\big((\Pi^2_{\underline{\lambda}})^W, \Gamma(\Theta)^W, T_{\underline{\lambda}}A(N)^W\big)$ where $W = T_{\underline{\lambda}}A(N) \cap \Pi^2_{\underline{\lambda}}$ and the superscript means everything is computed on the reduced subspace with respect to W.

Proof. In view of Proposition 4.4 and Remark 4.3 it only remains to prove that

$$\operatorname{ind}^- Q_{V^{\perp_Q}} = i\big(\Pi^2_{\underline{\lambda}}, \Gamma(\Theta), T_{\underline{\lambda}}A(N)\big).$$

Since $(v_0, v, v_1) \in V^{\perp_Q}$, we have that, $\forall w \in L^2[0,1], \forall \nu \in \Pi$:

$$\int_0^1 \left[H_t(v, w) + \sigma\left(Z_0 v_0 + \int_0^t Z_\tau v(\tau) d\tau, Z_t w(t)\right)\right] dt = \sigma\left(\int_0^1 Z_t w(t) dt, \nu\right).$$

Combining the last expression with (4.10) gives us:

$$\begin{aligned} Q(\hat{v}) &= -\sigma\left(Z_0 v_0 + \int_0^1 Z_t v(t) dt, Z_1 v_1\right) - \sigma\left(\int_0^1 Z_t v(t) dt, \nu\right) \\ &= -\sigma(\xi, Z_1 v_1) + \sigma(Z_1 v_1 + Z_0 v_0, \nu) \\ &= -\sigma(\nu, Z_0 v_0) + \sigma(\xi + \nu, -Z_1 v_1) \end{aligned} \tag{4.15}$$

where we have used that

$$\xi = Z_0 v_0 + Z_1 v_1 + \int_0^1 Z_t v(t) dt \in \Pi.$$

From the derivation of the Jacobi equation it follows that

$$\eta(0) = \nu + Z_0 v_0 \quad \eta(1) = \nu + Z_0 v_0 + \int_0^1 Z_t v(t) dt = \nu + \xi - Z_1 v_1.$$

Hence the restriction of Q to $V^{\perp Q}$ coincides with the quadratic form

$$m\big(\Pi^2, \Gamma(\Phi), T_{\lambda(0)}A(N_0) \times \tilde{\Phi}_*^{-1} T_{\lambda(1)}A(N_1)\big).$$

Note that $Q|_{V^{\perp Q}}$ is actually defined on a slightly smaller space, because $Z_0 v_0$ do not span the whole $T_{\lambda(0)}A(N_0)$ and similarly $Z_1 v_1$ does not span $\tilde{\Phi}_*^{-1} T_{\lambda(1)}A(N_1)$ correspondingly.

Nevertheless we obtain the correct Maslov form. In fact, the map $Z_0 : \mathbb{R}^k \to T_{\lambda(0)}A(N_0)$ is injective and its image is transversal to $\Pi \cap T_{\lambda(0)}A(N_0)$ (and the same is true for the Z_1). The sum of spaces $\Pi \cap T_{\lambda(0)}A(N_0)$ and $\Pi \cap \tilde{\Phi}_*^{-1} T_{\lambda(1)}A(N_1)$ lies in the kernel of the Maslov form. Removing it does not change the domain since $\operatorname{Im} Z_0 + \Pi = T_{\lambda(0)}A(N_0) + \Pi$ (and similarly for Z_1).

Hence we can either reduce by $\Pi \cap T_{\lambda(0)}A(N_0) \oplus \Pi \cap \tilde{\Phi}_*^{-1} T_{\lambda(1)}A(N_1)$ or work on the original space. The index is the same.

We now apply the map $I \times (\tilde{\Phi}_t)_*$ to each Lagrangian space inside the Maslov index of the triple above. By Remark 4.3 we get

$$i\big(\Pi^2, \Gamma(\Phi), T_{\lambda(0)}A(N_0) \times \tilde{\Phi}_*^{-1} T_{\lambda(1)}A(N_1)\big) = i\big(\Pi^2_{\underline{\lambda}}, \Gamma(\Theta), T_{\underline{\lambda}}A(N)\big).$$

□

4.4.2 Proof of Theorem 1.1

Before proving the general formula, we prove a corollary of Proposition 4.5. Assume that we have an optimal control problem as in eq. (B.2) and two sets of possible boundary conditions:

$$(q(0), q(1)) \in N_0 \times N_1 =: N$$

and

$$(q(0), q(1)) \in \tilde{N}_0 \times \tilde{N}_1 =: \tilde{N}.$$

and assume that a curve $\lambda : [0,1] \to TM$ is an extremal in both problems simultaneously, which simply means that λ is a solution of the Hamiltonian system of PMP and satisfies the transversality conditions for both boundary conditions at the same time, i.e. λ_i annihilates the sum $T_{\lambda_i}N_i + T_{\lambda_i}\tilde{N}_i$. A relevant example to keep in mind is when $N \subset \tilde{N}$. In this case if λ satisfies the transversality conditions for $\tilde{N}$ it satisfies the transversality conditions for N automatically.

Consider the two second variations Q_N and $Q_{\tilde{N}}$ corresponding to the two optimal control problems with boundary conditions like above. Using Proposition 4.5 we can find the difference between the Morse indices of those two quadratic forms.

Corollary 4.1. *Using the notations of this section the following formula holds*

$$\begin{aligned} \operatorname{ind}^- Q_{\tilde{N}} - \operatorname{ind}^- Q_N = {} & i\big(T_{\underline{\lambda}}A(N), \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big) + \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N)) + \\ & - \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N})) + \dim(T_{\pi(\underline{\lambda})}N \cap T_{\pi(\underline{\lambda})}\tilde{N}) + \\ & - \dim T_{\pi(\underline{\lambda})}N. \end{aligned} \tag{4.16}$$

Proof. Apply Proposition 4.5 to get an expression for $\operatorname{ind}^- Q_{\tilde{N}}$ and $\operatorname{ind}^- Q_N$. Subtracting one from the other gives

$$\begin{aligned} \operatorname{ind}^- Q_{\tilde{N}} - \operatorname{ind}^- Q_N = {} & i\big(\Pi^2_{\underline{\lambda}}, \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big) - i\big(\Pi_{\underline{\lambda}}, \Gamma(\Theta), T_{\underline{\lambda}}A(N)\big) + \\ & + \dim(\Gamma(\Theta) \cap \Pi_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(N)) - \dim(\Gamma(\Theta) \cap \Pi_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(\tilde{N})) \end{aligned}$$

Apply formula (A.7) with $L_0 = \Pi^2_{\underline{\lambda}}$, $L_1 = \Gamma(\Theta)$, $L_2 = T_{\underline{\lambda}}A(\tilde{N})$ and $L_3 = T_{\underline{\lambda}}A(N)$. After cancellations this results in

$$\begin{aligned}\operatorname{ind}^- Q_{\tilde{N}} - \operatorname{ind}^- Q_N = {}& i\big(\Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N}), T_{\underline{\lambda}}A(N)\big) - i\big(\Pi^2_{\underline{\lambda}}, T_{\underline{\lambda}}A(\tilde{N}), T_{\underline{\lambda}}A(N)\big) + \\ & + \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N)) - \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N})) - \\ & - \dim(\Pi^2_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(\tilde{N})) + \dim(\Pi^2_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N})).\end{aligned}$$

Terms $\dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N))$, $\dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N}))$ are already exactly as in the formula of the statement. It remains to simplify all of the remaining terms.

By formula (A.8)

$$i\big(\Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N}), T_{\underline{\lambda}}A(N)\big) = i\big(T_{\underline{\lambda}}A(N), \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big)$$

since we have an even permutation of subspaces inside. By lemma A.3 we have

$$i\big(\Pi^2_{\underline{\lambda}}, T_{\underline{\lambda}}A(\tilde{N}), T_{\underline{\lambda}}A(N)\big) = i\big(T_{\underline{\lambda}}A(N), \Pi^2_{\underline{\lambda}}, T_{\underline{\lambda}}A(\tilde{N})\big) = 0.$$

Finally, straight from the definition of an annihilator, it follows that

$$\dim(\Pi^2_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(\tilde{N})) = 2\dim M - \dim T_{\pi(\underline{\lambda})}\tilde{N}$$

and

$$\dim(\Pi^2_{\underline{\lambda}} \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N})) = 2\dim M - \dim T_{\pi(\underline{\lambda})}\tilde{N} - \dim T_{\pi(\underline{\lambda})}N + \dim(T_{\pi(\underline{\lambda})}N \cap T_{\pi(\underline{\lambda})}\tilde{N}).$$

Combining all of the above results in formula (4.16).

□

Remark 4.4. Notice that if $N = \{q_0\} \times \{q_1\}$ we obtain exactly the formula from Proposition 4.4 as expected. Another necessary remark is that formula (4.16) might seem asymmetric at first. We expect, that if we exchange N and $\tilde{N}$, then the resulting right-hand side will change sign. This is not entirely obvious just from the expression itself. However, this is indeed the case, because the difference between $i\big(T_{\underline{\lambda}}A(N), \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big)$ and $i\big(T_{\underline{\lambda}}A(\tilde{N}), \Gamma(\Theta), T_{\underline{\lambda}}A(N)\big)$ is not zero, but an expression involving dimensions of intersections of various subspaces as can be seen from formula (A.7).

Now we are ready to prove Theorem 1.1. We will reduce the case of general boundary conditions $(q_0, q_1) \in N \subseteq M \times M$ to the case with separated boundary conditions by introducing extra dummy variables.

Proof of Theorem 1.1. Consider optimal control problem as in eq. (B.2). We can lift it to an optimal problem on $M \times M$ by considering a new control system:

$$\begin{cases} \dot{x} = 0, \\ \dot{q} = f^t_{u(t)}(q), \end{cases} \tag{4.17}$$

with boundary conditions

$$(x(0), q(0), x(1), q(1)) \in \Delta \times N \subset M^4. \tag{4.18}$$

It is clear that there is a one-to-one correspondence between admissible curves of the original problem and admissible curve of (4.17)-(4.18). For this reason we can consider admissible curves (4.17)-(4.18) which minimize the functional of eq. (B.2). For clarity, assume that the maximum condition of PMP defines a regular function in a neighbourhood of the extremal. The Hamiltonian system of PMP is then given by

$$\begin{cases} \dot{\mu} = 0, \\ \dot{\lambda} = \vec{H}^t(\lambda), \end{cases} \qquad \lambda, \mu \in T^*M.$$

and its flow is given by $I \times \Psi_t$. This is not restrictive since in the non regular case the linearisation of this flow will be simply replaced by $\tilde{\Phi}_*\Phi$, as already mentioned several times.

We can now apply directly Corollary 4.1 to the boundary conditions $\Delta \times \tilde{N}$ and $\Delta \times N$. In order to see that everything indeed reduces to formula (1.3) without writing explicitly the lengthy formula here, let us go term by term starting from the last one. Let $\underline{\lambda} = (\lambda(0), \lambda(0), \lambda(0), \lambda(1))$. We have

$$\begin{aligned} \dim\left(T_{\pi(\underline{\lambda})}(\Delta \times N) \cap T_{\pi(\underline{\lambda})}(\Delta \times \tilde{N})\right) - \dim\left(T_{\pi(\underline{\lambda})}(\Delta \times N)\right) = \\ = \dim T_{\pi(\underline{\lambda})}\Delta + \dim(T_{\pi(\underline{\lambda})}N \cap T_{\pi(\underline{\lambda})}\tilde{N}) - \dim T_{\pi(\underline{\lambda})}\Delta - \dim T_{\pi(\underline{\lambda})}N \\ = \dim(T_{\pi(\underline{\lambda})}N \cap T_{\pi(\underline{\lambda})}\tilde{N}) - \dim T_{\pi(\underline{\lambda})}N. \end{aligned}$$

In order to deal with the last term choose Darboux coordinates around $\lambda(0)$ and $\lambda(1)$ which fix the horizontal and the vertical subspaces. For this computation we identify $T_{\lambda(0)}(T^*M) \simeq T_{\lambda(1)}(T^*M) =: \Sigma$. Let $S : \Sigma \to \Sigma$ be the map, which changes the sign of the vertical part. In Darboux coordinates it is given by the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In particular, we have

$$\sigma(S\mu_1, S\mu_2) = -\sigma(\mu_1, \mu_2), \qquad \forall \mu_1, \mu_2 \in \Sigma.$$

Let us write down explicitly each individual subspace entering the formula. Due to our conventions of signs for the symplectic form on $(T^*M)^4$ we have $-\sigma$ on the first two copies of T^*M and σ on the last two. We have to distinguish the two definitions of annihilator here, $\hat{A}(N)$ is the annihilator when we use $\sigma \oplus \sigma$ whereas $A(N)$ is given by eq. (A.10), and is the right object to use when the symplectic form is $(-\sigma) \oplus \sigma$. Notice once again that $S \times I(\hat{A}(N)) = A(N)$.

$$\begin{aligned} T_{\underline{\lambda}}A(\Delta \times N) &= \{(S\xi, \xi, \nu_1, \nu_2) : \xi \in \Sigma, (\nu_1, \nu_2) \in T_{\underline{\lambda}}\hat{A}(N)\}, \\ T_{\underline{\lambda}}A(\Delta \times \tilde{N}) &= \{(S\tilde{\xi}, \tilde{\xi}, \tilde{\nu}_1, \tilde{\nu}_2) : \tilde{\xi} \in \Sigma, (\tilde{\nu}_1, \tilde{\nu}_2) \in T_{\underline{\lambda}}\hat{A}(\tilde{N})\}, \\ \Gamma(I \times \Theta) &= \{(\eta_1, \eta_2, \eta_1, \Theta\eta_2) : \eta_1, \eta_2 \in \Sigma\}. \end{aligned} \tag{4.19}$$

From expressions in (4.19) it directly follows that

$$\begin{aligned} \dim(\Gamma(I \times \Theta) \cap T_{\underline{\lambda}}A(\Delta \times N)) = \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N)) \\ \dim(\Gamma(I \times \Theta) \cap T_{\underline{\lambda}}A(\Delta \times N) \cap T_{\underline{\lambda}}A(\Delta \times \tilde{N})) = \dim(\Gamma(\Theta) \cap T_{\underline{\lambda}}A(N) \cap T_{\underline{\lambda}}A(\tilde{N})). \end{aligned}$$

In order to simplify the Maslov index term, we note that the intersection of annihilators contains the following isotropic subspace

$$W = \{(S\xi, \xi, 0, 0) : \xi \in \Sigma\}.$$

For this reason we can perform a reduction to the space $W^{\perp}/W$. We have

$$W^{\perp} = \{(S\xi_2, \xi_2, \xi_3, \xi_4) : \xi_i \in \Sigma\}.$$

Thus we can identify $W^{\perp}/W$ with the image of the projection $\pi_1 : \Sigma^4 \to \Sigma^2$ to the last two terms $(\Sigma \times \Sigma, \sigma \oplus \sigma)$. Let us consider the space $(W^{\perp} + W) \cap \Gamma(I \times \Theta)$, it is straight forward to check that its projection is the subspace $\{(\eta, \Theta\eta) : \eta \in \Sigma\}$. We can now calculate the Maslov form on the reduced space. First of all we write down the equation defining the subspace:

$$\begin{cases} \nu_1 + \tilde{\nu}_1 = \eta, \\ \nu_2 + \tilde{\nu}_2 = \Theta\eta. \end{cases}$$

where $(\nu_1, \nu_2) \in T_{\underline{\lambda}}\hat{A}(N)$, $(\tilde{\nu}_1, \tilde{\nu}_2) \in T_{\underline{\lambda}}\hat{A}(\tilde{N})$, $\eta \in \Sigma$. It follows that

$$m\left((\eta, \Theta\eta)\right) = \sigma(\nu_1, \tilde{\nu}_1) + \sigma(\nu_2, \tilde{\nu}_2) = -\sigma(S\nu_1, S\tilde{\nu}_1) + \sigma(\nu_2, \tilde{\nu}_2),$$

But this is exactly the Maslov form $m\big(T_{\underline{\lambda}}A(N), \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big)$. Hence

$$i\big(T_{\underline{\lambda}}A(\Delta \times N), \Gamma(I \times \Theta), T_{\underline{\lambda}}A(\Delta \times \tilde{N})\big) = i\big(T_{\underline{\lambda}}A(N), \Gamma(\Theta), T_{\underline{\lambda}}A(\tilde{N})\big).$$

□

Chapter 5

Determinant of the Second Variation, Hill-type formulae and stability

The purpose of this chapter is to give some formulas to compute the determinant of the Second Variation for a general optimal control problem. The main result (Theorem 1.2) provides an extension of the classical Hill's formula in celestial mechanics (first appearing in [35]) and proved by Poincaré ([52]) to a quite general class of optimal control problems. Classical Hill's formula provides a way to compute the Fredholm determinant of an operator of the form $1+K$ where K is a trace class operator in terms of the fundamental solution of a linear Hamiltonian system of the form $\dot{x} = JH_t x$.

As discussed in Section 1.3 and Chapter 3, however, the second variation will not in general be a trace class operator. It is possible, though, to produce certain finite dimensional approximation of the Second Variation for which the *partial* trace and determinant converge. This strategy is explained and applied for example in [39] or [20] using Poincaré's original argument. However, if the operator is not trace class, this limits may very well depend on the approximation chosen. Using theorem 1.5 and [7][Theorem 1] we know that, at least under the analyticity assumption on H_t and Z_t, that the following limits exists and are finite:

$$\lim_{\epsilon\to 0^+} \sum_{\lambda\in \mathrm{Spec}\,(K), |\lambda|>\epsilon} m(\lambda)\lambda, \qquad \lim_{\epsilon\to 0^+} \prod_{\lambda\in \mathrm{Spec}\,(K), |\lambda|>\epsilon} (1+\lambda)^{m(\lambda)}.$$

It is worth noticing that this limits correspond to a different choice of finite dimensional approximation from the one in [20, 39]. Thus our formulas have different normalizations. The result we are going to prove here is an equality, for $s\in\mathbb{R}$, of the form:

$$\det(1+sK) := \lim_{\epsilon\to 0^+} \prod_{\lambda\in \mathrm{Spec}\,(K), |\lambda|>\epsilon} (1+s\lambda)^{m(\lambda)} = (-1)^{\dim(M)} a e^{bs} \det(M_1^s + M_2^s \tilde{\Phi}_* \Phi^s). \tag{5.1}$$

where M_1^s and M_2^s are suitable matrices which depend on the boundary conditions we are imposing; $a>0$ and b are explicit constants and $\tilde{\Phi}_*$ the flow we use to trivialize a neighbourhood of the extremal (see eq. (B.11)). The matrix Φ^s is the fundamental solution of a Jacoby-type equation defined in eq. (5.15). For $s=1$ it coincides with the already mentioned standard one given in eq. (1.2). As remarked in Chapter 4, $\tilde{\Phi}_*\Phi^1$ is the linearisation of the maximized Hamiltonian whenever the latter is defined.

This kind of formulas have attracted a lot of interest since their appearance, especially when dealing with periodic minimizers. The main reason is that they provide a connection between the stability properties of the trajectories and the parity of their Morse index. See for example [20] for a detailed account about Hill's formula for both continuous and discrete Lagrangian system with quasi periodic conditions or [40, 42] for the case of more general boundary conditions.

As far as the connection with stability goes, in [20] for instance, using Hill's formula, some criteria for the *linear instability* of closed geodesics are given. Similar results can be found in [56, 51] or [57, 58, 42] where one deals with periodic extremal with additional symmetries.

Let us explain briefly the idea behind some of these results. It is well known, see for example [49, 27], that the spectrum of a symplectic matrix is stable with respect to inversion and conjugation. This means that if $\Phi \in \mathrm{Sp}(2n)$ and $\lambda \in \mathrm{Spec}\,(\Phi)$ then $\bar{\lambda}$, λ^{-1} and $\bar{\lambda}^{-1}$ belong to the spectrum of Φ too. The orbit of any point $z \in \mathbb{C}$ with respect to inversion and conjugation is made up of either 4, 2 or 1 point. If $\Im(z) \neq 0$ and z does not lie on $\mathbb{S}^1$, it contains four points. If it lies on $\mathbb{S}^1 \cup \mathbb{R} \setminus \{\pm 1\}$ it contains 2 and if it z is ± 1 or 0 it is a fixed point. Thus a fundamental for this action is the set D defined as:

$$D = D_1 \cup D_2 \cup D_3, \quad D_1 = \{z : \Im(z) > 0, |z| > 1\},$$
$$D_2 = \{z : |z| = 1, \Im(z) > 0\} \cup \{z : Im(z) = 0, |z| > 1\}, \quad D_3 = \{\pm 1, 0\}.$$

Denote by $m(\lambda)$ the algebraic multiplicity of an eigenvalue λ of Φ. If 1 is not in the spectrum of Φ the determinant of $\Phi - 1$ can be decomposed as follows:

$$\det(\Phi - 1) = 2^{m(-1)} \prod_{\lambda \in D_1 \cap \mathrm{Spec}\,\Phi} \left(\frac{((\lambda-1)(\bar{\lambda}-1))^2}{|\lambda|^2} \right)^{m(\lambda)} \prod_{\lambda \in D_2 \cap \mathrm{Spec}\,\Phi} \left(\frac{-(\lambda-1)^2}{\lambda} \right)^{m(\lambda)}$$

Clearly eigenvalues in D_1 just give a positive contribution to the determinant. It is straightforward to check that if λ lives in $D_2 \cap \mathbb{S}^1$ then $\frac{-(\lambda-1)^2}{\lambda} = |\lambda - 1|^2$, which is positive, and that if $\lambda \in D_2 \cap \mathbb{R}_-$ the contribution is positive too. It follows that the sign of the determinant is determined by the presence of an odd number of *real positive* eigenvalues greater than 1, counted with algebraic multiplicity.

On the other hand the sign of the determinant of the Second Variation is determined by the number of negative eigenvalues, so assuming that it is non degenerate $\mathrm{sg}(\det(Q)) = (-1)^{\mathrm{ind}(Q)}$. Now, Theorem 5.1 and Theorem 5.2 link the determinant of fundamental solution of Jacobi equation (eq. (1.2)) and the determinant of the Second Variation through the following formula:

$$\det(Q)(-1)^{\dim(M)} = a \det(\Phi - 1), \quad a > 0.$$

Which in particular implies that non degenerate extremals with odd index on even dimensional manifolds (and extremal with even index on odd dimensional manifolds) are linearly unstable.

However, if the critical point is degenerate, as often happens for periodic geodesics, Hill's formula alone is of little help since both sides of the equation are zero. Two solutions to this problem have been proposed in [20]. The first one consist in removing the known first integrals reducing the system to new one, on a smaller space. Then proving a Hill formula for the reduced system and finding a relation for the Morse index of the starting trajectory and the reduced one. The second is via a perturbative argument. One changes the domain of the quadratic form $W^{1,2}([0,T], \mathbb{C}^m)$ with periodic boundary conditions to

$W^{1,2}([0,T],\mathbb{C}^m)$ with $v(0) = \rho v(T)$ and $\rho \in \mathbb{S}^1$. Then the sign of its determinant is the same as $(-1)^{\dim(M)}\rho^{-\dim(M)}\det(\Phi-\rho)$. The same argument explained above with 1 replaced by ρ shows that this is negative if and only if there is an odd number of real eigenvalues grater then 1 (counted with algebraic multiplicity). In the proof we already construct a perturbation of the second variation, namely the characteristic polynomial in eq. (5.1). It will be object of further investigation whether one can exploit this natural parameter to give some new stability criteria.

A by-product of our discussion will be integral expressions for the trace of the compact part of the second variation (see for instance lemma 5.3). Similar formulas have been employed in [36, 42] to get bounds on the Morse index of an extremal, its degeneracy and stability properties. These ideas trace back to the work of Krein [59, 47, 48]. He considered equations of type $\dot{x} = \lambda J H_t x$ with $H_t \geq 0$ and $\int_0^T H_t dt > 0$ and found a bound on $|\lambda|$, in terms of H_t, which guaranties that the solution is spectrally stable. Here by spectral stability we mean that the eigenvalues of the fundamental matrix lie on the unit circle. The arguments rely heavily on the positivity assumptions, however in [41] the authors managed to prove some inequalities on the number of eigenvalues on the unit circle for stable systems subjects to positive perturbations. All this criteria are express through bounds on the trace of the compact part of the second variation.

Another application of all those formulas is the following observation. Take a summable sequence a_n of positive real numbers and consider the sequence $S_k := \sum_n a_n^k$. Then, for k sufficiently large, it is monotone. Denote by $L := \lim_{k\to\infty} S_k$. It is straightforward to see that $L \in \mathbb{N}\cap\{0,+\infty\}$. In fact if all a_n are smaller than 1 $L = 0$, L is finite and non zero if and only if there are L a_n equals to 1 and $+\infty$ if there is at least one $a_n \geq 1$. This in particular shows that if $\sum_{\lambda\in\mathrm{Spec}\,(K)} \lambda^{2k}$ is smaller than 1 the extremal in consideration is a non degenerate minimum and if $\lim_k S_{2k} = n_0$ its Morse index is zero with at most n_0 degeneracy. In this case one can read off the nullity evaluating $\lim_k S_{2k+1}$. Unfortunately the expression for the trace we get are not so simple and same work is still needed to understand how much information one can really extract form this relation. This topic too will be object of further investigations.

In Section 5.2 we are going to present some application and specification of our result. We deduce the classical Hill's formula (Theorem 5.1) and a slightly more general version for non mechanical systems on $\mathbb{R}^n$ (Theorem 5.2). The rest of the chapter is devoted to the proof of Theorem 1.2. Much of the material presented here is contained in [17].

5.1 The Second Variation

The aim of this section is to recall the definition of Second Variation we are going to use and introduce a proper metric structure on the space of variations to compute the eigenvalues. Details for the first part are given in Section 2.3. Let $n_0, n_1 \in \mathbb{N}$ and consider the Hilbert space $\mathcal{H} = \mathbb{R}^{n_0} \oplus L^2([0,1],\mathbb{R}^k) \oplus \mathbb{R}^{n_1}$. The scalar products that we will employ are given by the direct sum of scalar products on the summands. Let (Σ,σ) be a symplectic space and consider a linear map $Z : \mathcal{H} \to \Sigma$ given by:

$$Z(u) = Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1, \quad u = (u_0, u_t, u_1) \in \mathcal{H}.$$

Suppose that $\Pi \subset \Sigma$ is a Lagrangian subspace transversal to the image of the map Z and define $\mathcal{V} = Z^{-1}(\Pi)$. For an appropriate choice of Z and Π, the *second variation* will be the quadratic form given in the following definition:

Definition 1 (Second Variation)**.** The second variation at $\tilde{u}$ is the quadratic form defined on $\mathcal{V} \subseteq \mathcal{H}$:

$$Q(u) = \int_0^1 \left(\langle u_t, u_t \rangle + \sigma(Z_t u_t, \int_0^t Z_\tau u_\tau d\tau + Z_0 u_0) \right) dt + \\ + \sigma(Z_0 u_0 + \int_0^1 Z_t u_t dt, Z_1 u_1).$$

Moreover we will make the following assumptions: the extremal of the original problem is *strictly normal* and satisfies *Legendre strong condition*, this means that for $t \in [0,1]$:

$$X_t := \pi_* Z_t \text{ satisfies } \int_0^1 X_t^* X_t > 0, \quad -H_t > \alpha > 0.$$

5.1.1 The scalar product on the space of variations

As already mentioned, we will assume through out this chapter *Legendre strong condition*, that is the matrix $-H_t$ is positive definite on $[0,1]$, with uniformly bounded inverse. This allows to use $-H_t$ to define an Hilbert structure on $L^2([0,1], \mathbb{R}^k)$ equivalent to the standard one. We have still to define the scalar product on a subspace transversal to $\mathcal{V}_0 = \{u_0 = u_1 = 0\}$. A natural choice would be to introduce two metrics on $T_{\lambda_0}T^*M$ and $T_{\lambda_1}T^*M$ and pull them back to the space of controls using the maps Z_0 and $\tilde{\Phi}_* Z_1 : \mathbb{R}^n \to T_{\lambda_0}T^*M$. Let us call any such metrics g_0 and g_1.

Definition 5.1. For any $u, v \in \mathcal{H}$ define:

$$\langle u, v \rangle = -\int_0^1 H_t(u_t, v_t) dt + g_0(Z_0 u_0, Z_0 v_0) + g_1(\tilde{\Phi}_* Z_1 u_1, \tilde{\Phi}_* Z_1 v_1)$$

Since the symplectic form σ is a skew-symmetric bilinear form there exists a g_i-skew-symmetric linear operator J_i such that:

$$g_i(J_i X_1, X_2) = \sigma(X_1, X_2), \quad \forall X_1, X_2 \in T_{\lambda_i}T^*M, \, i = 0, 1.$$

Notice that in general the matrix J_i fails to be an almost complex structure and $J_i^2 \neq -1$, in other words the metric and symplectic structures need not be compatible.

In terms of the symplectic form the scalar product can be written as:

$$\langle u, v \rangle = -\int_0^1 H_t(u_t, v_t) dt + \sigma(J_0^{-1} Z_0 u_0, Z_0 u_0) + \sigma(J_1^{-1} \tilde{\Phi}_* Z_1 u_1, \tilde{\Phi}_* Z_1 u_1)$$

Now we are going to compute the orthogonal complement to $\mathcal{V}$ inside $\mathcal{H}$ using the Hilbert structure just introduced. We will denote by the symbol $\perp_i$ the orthogonal complement, in $T_{\lambda_i}T^*M$, with respect to the scalar product g_i.

Lemma 5.1. *With this choice of scalar product the orthogonal complement to $\mathcal{V}$ is given by:*

$$\mathcal{V}^\perp = \{(v_0, -H_t^{-1} J Z_t^* \nu, v_1) : \nu \in \Pi\}$$

where v_0 and v_1 are determined by the following conditions:

$$Z_0 v_0 - J_0 \nu \in \operatorname{Im} Z_0^{\perp_0}, \quad \tilde{\Phi}_* Z_1 v_1 - J_1 \tilde{\Phi}_* \nu \in \operatorname{Im} \tilde{\Phi}_* Z_1^{\perp_1}.$$

Proof. Suppose that $u \in \mathcal{V}^{\perp}$. Let us test v against variations with fixed end-points. Namely such that $u_i = 0$ and $\int_0^1 Z_t u_t dt \in \Pi$. Recall that Π is Lagrangian, thus the condition $\int_0^1 Z_t u_t dt \in \Pi$ can be equivalently formulated as $\sigma(\int_0^1 Z_t u_t dt, \nu) = 0$ for all $\nu \in \Pi$. This implies that the linear functional defining the subspace $\{u : u_i = 0, \int_0^1 Z_t u_t dt \in \Pi\}$ are $u_t \mapsto \int_0^1 \sigma(Z_t u_t, \nu) dt$ and so:

$$\langle v, u \rangle = -\int_0^1 \langle H_t v_t, u_t \rangle dt = 0, \forall u \in V \iff v_t = -H_t^{-1} \sigma(\nu, Z_t \cdot), \quad \nu \in \Pi.$$

Thus it follows that $v_t = -H_t^{-1} Z_t^* J \nu$. Now take a generic $u \in \mathcal{V}$ and compute:

$$\langle v, u \rangle = \sigma(\nu, \int_0^1 Z_t u_t dt) + \sigma(J_0^{-1} Z_0 v_0, Z_0 u_0) + \sigma(J_1^{-1} \tilde{\Phi}_* Z_1 v_1, \tilde{\Phi}_* Z_1 u_1)$$

Now we get rid of the term $\int_0^1 Z_t u_t dt$ using the fact that $u \in \mathcal{V}$ and thus:

$$\sigma\left(\nu, \int_0^1 Z_t u_t dt\right) = -\sigma(\nu, Z_0 u_0 + Z_1 u_1)$$

It follows that

$$\begin{aligned} \langle v, u \rangle &= -\sigma(\nu, Z_0 u_0 + Z_1 u_1) + \sigma(J_0^{-1} Z_0 v_0, Z_0 u_0) + \sigma(J_1^{-1} \tilde{\Phi}_* Z_1 v_1, \tilde{\Phi}_* Z_1 u_1) \\ &= \sigma(J_0^{-1} Z_0 v_0 - \nu, Z_0 u_0) + \sigma(J_1^{-1} \tilde{\Phi}_* Z_1 v_1 - \tilde{\Phi}_* \nu, \tilde{\Phi}_* Z_1 u_1) \end{aligned}$$

Notice that the two quantities are independent. If $\langle u, v \rangle = 0$ then both $\sigma(J_0^{-1} Z_0 v_0 - \nu, Z_0 u_0)$ and $\sigma(J_1^{-1} \tilde{\Phi}_* Z_1 v_1 - \tilde{\Phi}_* \nu, \tilde{\Phi}_* Z_1 u_1)$ must be zero at the same time.

This follows from the fact that the map $u \mapsto \pi_* \int_0^1 Z_t u_t dt$ is surjective. You can build infinitesimal variations of the form $(u_0, u_t, 0)$ and $(0, u_t, u_1)$ in such a way that $Z_0 u_0$ and $Z_1 u_1$ span the whole $\mathrm{Im}(Z_0)$ and $\mathrm{Im}(Z_1)$ respectively.

This implies that v_0 and v_1 are completely determined by the value of ν and so:

$$\begin{aligned} \sigma(J_0^{-1} Z_0 v_0 - \nu, Z_0 u_0) &\iff Z_0 v_0 - J_0 \nu \in \mathrm{Im}(Z_0)^{\perp_0} \\ \sigma(J_1^{-1} \tilde{\Phi}_* Z_1 v_1 - \tilde{\Phi}_* \nu, Z_1 u_1) &\iff \tilde{\Phi}_* Z_1 v_1 - J_1 \tilde{\Phi}_* \nu \in \mathrm{Im}(\tilde{\Phi}_* Z_1)^{\perp_1} \end{aligned}$$

□

Now we use the Hilbert structure just introduced to write the quadratic form associated to compact part K of the second variation given in eq. (2.7).

Preliminarily we can perform the change of coordinates in L^2 sending $v_t \mapsto (-H_t)^{\frac{1}{2}} v_t$ and substituting Z_t with $Z_t(-H_t)^{-\frac{1}{2}}$. In this way the Hilbert structure on the interval becomes the standard one.

We introduce a further piece of notation, call pr_0 (respectively pr_1) the orthogonal projection on $\mathrm{Im}(Z_0)$ (respectively $\mathrm{Im}(\tilde{\Phi}_* Z_1)$) with respect to scalar product g_0 (respectively g_1). Let L be a partial inverse to $\tilde{\Phi}_* Z_1$ i.e. a map $L : T_{\lambda_0} T^* M \to \mathbb{R}^n$ defined by the relation $L \tilde{\Phi}_* Z_1 v_1 = v_1$. Set:

$$\Lambda(u) = L\, pr_1\, J_1 \tilde{\Phi}_* \Big(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1 \Big) \tag{5.2}$$

Lemma 5.2. *The second variation, as a bilinear form, can be expressed as:* $Q(u,v) = \langle u + Ku, v\rangle$ *where* $u, v \in \mathcal{V}$ *and* K *is the operator defined by:*

$$Ku = \begin{pmatrix} -u_0 \\ -Z_t^* J\Big(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0\Big) \\ -u_1 - \Lambda(u) \end{pmatrix} \tag{5.3}$$

where $\Lambda(u)$ *is given above, in eq.* (5.2).

Proof. A quick manipulation of the expression involving the symplectic form in Definition 2.1 yields the following:

$$\begin{aligned}
\int_{-1}^{2}\int_{-1}^{t} \sigma(Z_\tau u_\tau, Z_t v_t) d\tau dt &= \int_{-1}^{0}\int_{-1}^{t} \sigma(Z_\tau u_\tau, Z_t v_t) d\tau dt + \int_{0}^{1}\int_{-1}^{t} \sigma(Z_\tau u_\tau, Z_t v_t) d\tau dt \\
&\quad + \int_{1}^{2}\int_{-1}^{t} \sigma(Z_\tau u_\tau, Z_t v_t) d\tau dt \\
&= \int_0^1 \sigma\left(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0, Z_t v_t\right) dt + \sigma\Big(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1, Z_1 v_1\Big) \\
= \int_0^1 \sigma\left(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0, Z_t v_t\right) dt &+ g_1\Big(J_1 \tilde{\Phi}_*\Big(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1\Big), \tilde{\Phi}_* Z_1 v_1\Big)
\end{aligned}$$

Recall that Z_t is constant on $[0,1]^c$. Moreover the maps Z_0 and Z_1 take values in isotropic subspaces. We use this fact to simplify the expression in the first line. Now, it is clear that in the last term:

$$g_1\Big(J_1 \tilde{\Phi}_*\Big(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1\Big), \tilde{\Phi}_* Z_1 v_1\Big)$$

only the projection onto the image of $\tilde{\Phi}_* Z_1$ plays a role and it is straightforward to check that:

$$g_1\Big(J_1 \tilde{\Phi}_*\Big(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1\Big), \tilde{\Phi}_* Z_1 v_1\Big) = g_1(\tilde{\Phi}_* Z_1 \Lambda(u), \tilde{\Phi}_* Z_1 v_1).$$

Recall that we have normalize H_t to -1, thus the first summand can be rewritten as follows:

$$\int_0^1 \sigma\left(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0, Z_t v_t\right) dt = \int_0^1 \langle Z_t^* J\Big(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0\Big), v_t\rangle dt$$

If now we add and subtract $g(Z_0 u_0, Z_0 v_0)$ and $g(Z_1 u_1, Z_1 v_1)$ to single out the identity we obtain precisely the formula in the statement. □

5.2 Hill-type formulas

We present here some applications of theorem 1.2. We deduce Hill's formula for periodic trajectory and specify it to the eigenvalue problem for Schrödinger operators. In the second sub-section we present a variation of the classical Hill formula for systems with drift. In this section we will mainly deal with periodic and quasi-periodic boundary conditions, meaning that the boundary condition we take are of the form $N = \Gamma(f)$ for a diffeomorphism $f : M \to M$ of the state space.

5.2.1 Driftless systems and classical Hill's formula

In this section we consider drift-less systems with periodic boundary conditions on $\mathbb{R}^n$ and specify the formulas of Theorem 5.3 for this class of problems.

First of all let us explain what we mean by *drift-less* systems. Let $t \mapsto R_t$ a family of symmetric matrices of size $n \times n$ and let us denote by u a function in $L^\infty([0,1], \mathbb{R}^n)$.

Consider the following family of vector fields $f_u(q)$, their associated trajectories $q_u(t)$ and the action functional $\mathcal{A}(u)$:

$$\begin{aligned} &f_u(q) = u(t), \quad \begin{cases} \dot{q}_u = f_u(q) = u(t), \\ q(0) = q_0 \in \mathbb{R}^n \end{cases} \\ &\mathcal{A}(u) = \frac{1}{2}\int_0^1 |u|^2 - \langle R_t q_u(t), q_u(t)\rangle dt. \end{aligned} \tag{5.4}$$

We impose periodic boundary conditions, i.e. we take $N = \Delta = \{(q,q) \in \mathbb{R}^{2n} : q \in \mathbb{R}^n\}$. The Hamiltonian coming from the Maximum Principle takes the form:

$$H(p,q) = \max_{u \in \mathbb{R}^k} \langle p, u\rangle - \frac{1}{2}(|u|^2 - \langle R_t q, q\rangle) = \frac{1}{2}(\langle p, p\rangle + \langle R_t q, q\rangle). \tag{5.5}$$

Let us denote the flow generated by H by Ψ_t and fix a normal extremal λ_t and its control $\tilde{u}(t)$. The flow $\tilde{\Phi}$ we use to re-parametrize the space is given by the Hamiltonian:

$$\begin{aligned} &h_{\tilde{u}(t)}(p,q) = \langle p, \tilde{u}(t)\rangle + \frac{1}{2}\langle R_t q, q\rangle \quad \Rightarrow \quad \begin{cases} \dot{p} = -R_t q \\ \dot{q} = \tilde{u}(t). \end{cases} \\ &\tilde{\Phi}_t(p,q) = \begin{pmatrix} 1 & -\int_0^t R_\tau d\tau \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} + \begin{pmatrix} -\int_0^t \int_0^\tau R_\tau \tilde{u}(r) dr d\tau \\ \int_0^t \tilde{u}(\tau) d\tau \end{pmatrix} \\ &(\tilde{\Phi}_t)_* = \begin{pmatrix} 1 & -\int_0^t R_\tau d\tau \\ 0 & 1 \end{pmatrix}, \quad Z_t = (\tilde{\Phi}_t^{-1})_* \partial_u \vec{h}_u^t = \begin{pmatrix} \int_0^t R_\tau d\tau \\ 1 \end{pmatrix} \end{aligned}$$

Let us call $\hat{R} = -\int_0^1 R_\tau d\tau$. The annihilator to the diagonal is nothing else than the graph of the identity. We will now define Q^s as in eq. (5.18) and $\mathfrak{p}_Q(s)$ as $\det(Q^s)$ as in Section 1.2 (actually up to a scalar, but this is irrelevant). For $\eta \in T_{\lambda_0}T^*M$ set:

$$Q^s(\eta) = \begin{pmatrix} -1 & 1 \end{pmatrix}\begin{pmatrix} \eta \\ A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s \eta \end{pmatrix} = (A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s - 1)\eta \in T_{\lambda_0}T^*M.$$

Here the symplectic maps A_i^s for $i = 0,1$ are defined in section 1.2 and eq. (5.7). They are expressed in terms of the projection on the fibre Π_i and its orthogonal $\Pi_i^\perp$ and move, as $|s| \to +\infty$, any Lagrange subspace closer and closer to Π_i. A coordinate representation is given few lines below.

It is clear that the kernel of Q^s is precisely the intersection of the graph with the diagonal subspace. Since we are working on $\mathbb{R}^{2n}$ we can define the determinant of this map as:

$$\mathfrak{p}_Q(s) = \det(Q^s) = \det(A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s - 1)$$

As already mentioned in Section 1.2 this function is a multiple of the *characteristic polynomial* of K. It satisfies:

$$\mathfrak{p}_Q(s) = a e^{bs} \det(1 + sK), \quad a \in \mathbb{C}^*, b \in \mathbb{C}$$

Now we are going to compute the normalization factors. This is done essentially evaluating $\mathfrak{p}_Q(s)$ and its derivative in zero. This will give us the relations:

$$\mathfrak{p}_Q(0) = a, \quad a(b + \mathrm{tr}(K)) = \partial_s \mathfrak{p}_Q(s)|_{s=0}.$$

We have now to work a bit and write down precisely all the quantities appearing in the formulas. It is straightforward to compute the matrix representations of the maps A_0^s and A_1^s. In this setting the projections onto Π_0 and $\Pi_0^\perp$ are given by:

$$pr_{\Pi_0} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad pr_{\Pi_0^\perp} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

If we denote by G_0 and G_1 the restriction to $\Pi_0^\perp$ and $\Pi_1^\perp$ respectively, of the metrics chosen at the initial and final point, we see that:

$$A_0^s = \begin{pmatrix} 1 & (1-s)G_0 \\ 0 & 1 \end{pmatrix}, \quad A_1^s = \begin{pmatrix} 1 & (1-s)(G_1 - \hat{R}) \\ 0 & 1 \end{pmatrix}$$
$$A_1^s \tilde{\Phi}_* = \begin{pmatrix} 1 & s\hat{R} + (1-s)G_1 \\ 0 & 1 \end{pmatrix}$$

The value of Φ_t^s at $s = 0$ is given in Lemma 5.4, in this case, since $Y_t = \int_0^t R_s ds$ and $X_t = 1$, we obtain the following:

$$\Phi_t^0 = \begin{pmatrix} 1 & 0 \\ \Gamma & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$
$$\partial_s \Phi_t^s|_{s=0} = \begin{pmatrix} \Theta & 0 \\ \Omega & -\Theta^* \end{pmatrix} = \begin{pmatrix} \int_0^t \int_0^\tau R_s ds & 0 \\ \int_0^t \int_0^\tau \int_r^\tau R_s ds dr d\tau & -\int_0^t \int_0^\tau R_s ds \end{pmatrix}$$

We will still adopt the notation of Lemma 5.4 for the submatrices of Φ_t^0 and $\partial_s \Phi_t^s|_{s=0}$. Let us compute the value of $\mathfrak{p}_Q(s)$ in zero. Putting all together we have:

$$\mathfrak{p}_Q(s)|_{s=0} = \det\left(\begin{pmatrix} 1 & G_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \Gamma & 1 \end{pmatrix} \begin{pmatrix} 1 & G_0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

After a little bit of computation we find that $Q_s|_{s=0}$ satisfies:

$$Q_s|_{s=0} = \begin{pmatrix} G_1\Gamma & G_0 + G_1 + G_1\Gamma G_0 \\ \Gamma & \Gamma G_0 \end{pmatrix}, \quad \det(Q_s|_{s=0}) = (-1)^n \det(\Gamma)\det(G_1 + G_0),$$
$$(Q_s|_{s=0})^{-1} = \begin{pmatrix} -G_0(G_1+G_0)^{-1} & \Gamma^{-1} + G_0(G_1+G_0)^{-1}G_1 \\ (G_1+G_0)^{-1} & -(G_1+G_0)^{-1}G_1 \end{pmatrix}$$

We can compute the derivative $\det(Q_s)$ at $s = 0$, we find that:

$$\begin{aligned} \partial_s Q_s &= (\partial_s A_1^s)\tilde{\Phi}_* \Phi_1^s A_0^s + A_1^s \tilde{\Phi}_* (\partial_s \Phi_1^s) A_0^s + A_1^s \tilde{\Phi}_* \Phi_1^s (\partial_s A_0^s) \\ &= \begin{pmatrix} (G_1 - \hat{R})\Gamma & G_1 - \hat{R} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & G_0 \\ 0 & 1 \end{pmatrix} + \\ &\quad + \begin{pmatrix} 1 & G_1 \\ 0 & 1 \end{pmatrix} \left(\begin{pmatrix} \Theta & 0 \\ \Omega & -\Theta^* \end{pmatrix} \begin{pmatrix} 1 & G_0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & G_0 \\ 0 & \Gamma G_0 \end{pmatrix} \right) \end{aligned}$$

We use now Jacobi formula for the derivative of the determinant of a family of invertible matrices. It reads:

$$\partial_s \det(M_s) = \det(M_s)\,\mathrm{tr}(\partial_s M_s M_s^{-1}).$$

Without going into the detail of the actual computation, which at this point is just matrix multiplication, we have that:

$$\partial_s \det(Q_s)|_{s=0} = \mathrm{tr}(\partial_s(Q_s)Q_s^{-1})|_{s=0} = \mathrm{tr}((G_1+G_0)^{-1}(G_0-G_1+\hat{R})+\Gamma^{-1}\Omega)$$

The last quantity we have to compute is $\mathrm{tr}(K)$. To do so we use Lemma 5.3. Mind that in the statement of the Lemma one works with twice the variables, taking as state space $\mathbb{R}^n \times \mathbb{R}^n$ and using the symplectic form $(-\sigma)\oplus\sigma$ on $\mathbb{R}^{2n}\times\mathbb{R}^{2n}$. The quantities with $\tilde{}$ on top always refer to the system in $\mathbb{R}^{4n}$, where we have a trivial dynamic on the first factor and the boundary condition we impose are in this case $\Delta\times\Delta$ (see the beginning of Section 5.3.2 for more details). The formula given in the lemma reads:

$$\begin{aligned}\mathrm{tr}(K) = &-\dim(N) + \mathrm{tr}[\pi_*^1\tilde{\Phi}_*^{-1}\,pr_1\tilde{J}_1\tilde{\Phi}_*(\tilde{Z}_0)]\\ &+\mathrm{tr}\left[\Gamma^{-1}\left(\Omega+(\pi_*^2-\pi_*^1)\tilde{\Phi}_*^{-1}\,pr_1\tilde{J}_1\tilde{\Phi}_*(\int_0^1 \tilde{Z}_t Z_t^* J|_\Pi dt)\right)\right]\end{aligned}$$

Let us explain all the objects appearing in the formula. $\tilde{\pi}_*^i$ denotes the differential of the natural projection on the $i-$th factor. The matrix $\tilde{\Phi}_*$ is given here by:

$$\tilde{\Phi}_* = \begin{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} & 0\\ 0 & \begin{pmatrix}1 & R\\ 0 & 1\end{pmatrix}\end{pmatrix}$$

Moreover the matrices $\tilde{Z}_0$, $\tilde{Z}_t$ and $\tilde{Z}_1$ are:

$$\tilde{Z}_0 = \begin{pmatrix}0\\1\\0\\1\end{pmatrix},\quad \tilde{Z}_t = \begin{pmatrix}0\\0\\\int_0^t R_\tau d\tau\\1\end{pmatrix},\quad \tilde{Z}_1 = \begin{pmatrix}0\\1\\-R\\1\end{pmatrix},\quad \tilde{\Phi}_*\tilde{Z}_1 = \begin{pmatrix}0\\1\\0\\1\end{pmatrix}$$

The map pr_1 denotes the orthogonal projection onto the image of $\tilde{Z}_1$. We are using the scalar product $g_0\oplus g_1$ on $T_{\lambda_0}T^*M\times T_{\lambda_1}T^*M$ to define it. One can check that the following map is the coordinate representation of pr_1:

$$pr_1 = (G_0+G_1)^{-1}\begin{pmatrix}0&0&0&0\\0&G_0&0&G_1\\0&0&0&0\\0&G_0&0&G_1\end{pmatrix} \Rightarrow pr_1\tilde{J}_1 = (G_0+G_1)^{-1}\begin{pmatrix}0&0&0&0\\-1&0&1&0\\0&0&0&0\\-1&0&1&0\end{pmatrix}$$

Now everything reduces to some tedious matrix multiplications. The second term in the expression of the trace reads:

$$\mathrm{tr}[\pi_*^1\tilde{\Phi}_*^{-1}\,pr_1\tilde{J}_1\tilde{\Phi}_*(\tilde{Z}_0)] = \mathrm{tr}(\hat{R}(G_0+G_1)^{-1}).$$

For the third term notice that $(\pi_*^2-\pi_*^1)(\tilde{\Phi}_*)^{-1}pr_1$ is identically zero since $\tilde{\Phi}_*$ does not change the projection on the horizontal part and we are working with periodic boundary conditions. It follows we are left with $\mathrm{tr}(\Gamma^{-1}\Omega)$. Summing up we have computed that:

$$\mathrm{tr}(K) = \mathrm{tr}(\Omega\Gamma^{-1}+\hat{R}(G_0+G_1)^{-1}) - \dim(M).$$

We have thus reduce Theorem 1.2 to the following result:

Theorem 5.1 (Hill's formula)**.** *Let us consider a critical point of the functional given in eq. (5.4) with periodic boundary conditions. Let $1+K$ be the second variation, and Ψ the fundamental solution of $\dot{\Psi} = \vec{H}$, where H is given in eq. (5.5). Then the following equality holds:*

$$\det(1+K) = (-1)^n e^{-n} \det(G^{-1}) \det(1-\Psi)$$

where we chose $G_0 = G_1 = \frac{1}{2}G$.

Remark 5.1. If we are working on the interval $[0,T]$ instead of $[0,1]$ everything remains essentially unchanged. The only difference is that now $\det(\Gamma) = T^n$ and so this extra factor appears in the formulas.

We can apply the previous result to study boundary value problems for Sturm-Liouville equations. Let us illustrate the case of Schrödinger equation with periodic boundary conditions. Fix now the normal extremal of eq. (5.4) given by $(p(t), q(t)) = (0,0)$ and the relative control $\tilde{u} = 0$. Consider the cost $\tilde{R}_t = R_t + \lambda$, for $\lambda \in \mathbb{R}$. Consider the second variation of the functional

$$\mathcal{A}_\lambda(u) = \frac{1}{2}\int_0^1 |u_t|^2 + \langle (R_t+\lambda) q_u(t), q_u(t)\rangle dt$$

at the point $\tilde{u} = 0$. It is given by the operator $1 + K_\lambda$ where:

$$\langle K_\lambda(u), u\rangle = \lambda \left(\int_0^1 \int_0^t \langle (\tau - t) u(\tau), u(t)\rangle d\tau dt + \langle u_0, u_0\rangle \right) + \langle K_0(u), u\rangle$$

Then we have the following corollary:

Corollary 5.1. *Let $\lambda \in \mathbb{R}$, Ψ_λ the fundamental matrix of the lift to $\mathbb{R}^{2n}$ of the following ODE on $\mathbb{R}^{2n}$:*

$$\ddot{q}(t) = (R_t + \lambda) q(t)$$

The determinant of the operators $1+K_\lambda$ can be expressed as:

$$\det(1+K_\lambda) = (-1)^n e^{-n} \det(G^{-1}) \det(1-\Psi_\lambda),$$

where $G = 2G_0 = 2G_1$ as in the previous statement.

5.2.2 System with drift and Hill-type formulas

In this section we give a version of Hill's formula for linear systems with drift. They are again linear system with quadratic cost of the following form:

$$\begin{aligned} f_u(q) &= A_t q + B_t u(t), \qquad \begin{cases} \dot{q}_u = f_u(q) = A_t q_u + B_t u(t), \\ q(0) = q_0 \in \mathbb{R}^n \end{cases} \\ \mathcal{A}(u) &= \frac{1}{2}\int_0^1 |u|^2 + \langle R_t q_u(t), q_u(t)\rangle dt. \end{aligned} \tag{5.6}$$

Where A_t is $n \times n$ matrix and B_t a $n \times k$ one, both with possibly non-constant coefficients. The maximized Hamiltonian and the one used to reparametrize take the form:

$$H(p,q) = \langle p, A_t q\rangle + \frac{1}{2}\left(\langle B_t B_t^* p, p\rangle + \langle R_t q, q\rangle\right)$$

$$h^t_{\tilde{u}(t)} = \langle p, B_t \tilde{u}(t) + A_t q\rangle + \frac{1}{2}\langle R_t q, q\rangle.$$

In particular if we call $\hat{\Phi}_t$ the fundamental solution of $\dot{q} = A_t q$, we can write similar formulas as in the previous case using variation of constants. We find that the lift of $\hat{\Phi}_t$ to the cotangent bundle and its differential are:

$$\begin{aligned}
\tilde{\Phi}_t(p,q) &= \begin{pmatrix} (\hat{\Phi}_t^*)^{-1} & -(\hat{\Phi}_t^*)^{-1}\int_0^t \hat{\Phi}_\tau^* R_\tau \hat{\Phi}_\tau d\tau \\ 0 & \hat{\Phi}_t \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} \\
&+ \begin{pmatrix} -(\hat{\Phi}_t^*)^{-1}\int_0^t \hat{\Phi}_\tau^* R_\tau \hat{\Phi}_\tau \int_0^\tau \hat{\Phi}_r^{-1} B_r \tilde{u}(r) dr d\tau \\ \hat{\Phi}_t \int_0^t \hat{\Phi}_r^{-1} B_r \tilde{u}(r) dr \end{pmatrix} \\
(\tilde{\Phi}_t)_* &= \begin{pmatrix} (\hat{\Phi}_t^*)_*^{-1} & -(\hat{\Phi}_t^*)^{-1}\int_0^t \hat{\Phi}_\tau^* R_\tau \hat{\Phi}_\tau d\tau \\ 0 & \hat{\Phi}_t \end{pmatrix}
\end{aligned}$$

As boundary conditions manifold we take $N = \Gamma(\hat{\Phi}_t + \hat{\Phi}_t \int_0^t \hat{\Phi}_r^{-1} B_r \tilde{u}(r) dr)$. Notice that since only the tangent space matters the translation is irrelevant and it would be the same as if we considered $\Gamma(\hat{\Phi}_t)$ (or work with control $\tilde{u} = 0$). In particular we can take as T_0 and T_1 in the definition of Q_s (see eq. (5.18)) to be the following:

$$T_0^* J_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T_1^* J_1 = \begin{pmatrix} 0 & -(\hat{\Phi}_t)^{-1} \\ \hat{\Phi}_t^* & 0 \end{pmatrix}$$

Set as before $\hat{R} = -(\hat{\Phi}_t^*)^{-1}\int_0^t \hat{\Phi}_\tau^* R_\tau \hat{\Phi}_\tau d\tau$, the upper right minor of $(\tilde{\Phi}_t)_*$. A quick computation shows that:

$$T_1^* J_1 A_1^s \tilde{\Phi}_* = \begin{pmatrix} 1 & (1-s)\hat{\Phi}_t^* G_1 \hat{\Phi}_t + s\hat{\Phi}_t^* \hat{R} \\ 0 & 1 \end{pmatrix} := \hat{A}_1^s \tilde{\Phi}_t^*$$

In particular we are again brought to consider a function of the same type as the one in the previous section:

$$\mathfrak{p}_Q(s) = \det(\hat{A}_1^s \tilde{\Phi}_* \Phi_1^s A_0^s - 1).$$

In particular this means that the computation already done before are again valid, we ave just to substitute the new values of G_1 and Γ. It holds:

$$\begin{aligned}
\det(Q_s) &= (-1)^n \det(\Gamma)\det(\hat{\Phi}_t^* G_1 \hat{\Phi}_t + G_0), \\
\partial_s \det(Q_s)|_{s=0} &= \operatorname{tr}(((\hat{\Phi}_t^* G_1 \hat{\Phi}_t + G_0)^{-1}(G_0 - \hat{\Phi}_t^* G_1 \hat{\Phi}_t + \hat{\Phi}_t^* \hat{R}) + \Gamma^{-1}\Omega)
\end{aligned}$$

Now we have to apply Lemma 5.3 to compute the trace of the compact part of the second variation. Here pr_1 and $\tilde{Z}_1$ are different since we have changed boundary conditions. However we have the same kind of simplification as in the previous computation. Let us write explicitly the new objects:

$$Z_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \hat{\Phi}_t \end{pmatrix}, \quad pr_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & LG_0 & 0 & L\hat{\Phi}_t^* G_1 \\ 0 & 0 & 0 & 0 \\ 0 & \hat{\Phi}_t L G_0 & 0 & \hat{\Phi}_t L \hat{\Phi}_t^* G_1 \end{pmatrix}, \quad L = (G_0 + \hat{\Phi}_t^* G_1 \hat{\Phi}_t)^{-1}$$

In the end the trace reads:

$$\operatorname{tr}(K) = \operatorname{tr}(\Gamma^{-1}\Omega + \hat{\Phi}_t^* \hat{R}(\hat{\Phi}_t^* G_1 \hat{\Phi}_t + G_0)^{-1}) - \dim(M).$$

In particular, if we set $G = 2G_1$ and choose $G_0 = \hat{\Phi}_t^* G_1 \hat{\Phi}_t$ we have the following:

Theorem 5.2 (Hill's formula with drift)**.** *Suppose that a critical point of the functional given in eq. (5.6), with boundary conditions $N = \Gamma(\hat{\Phi}_t + \hat{\Phi}_t \int_0^t \hat{\Phi}_r^{-1} B_r \tilde{u}(r) dr)$, is fixed. Let $\underline{\Phi}_t$ be the lift to T^*M of $\hat{\Phi}_t$ given by the pullback of $1-$forms. Let G be our choice of scalar product, $\Gamma = \int_0^t \hat{\Phi}_\tau B_\tau B_\tau^* \hat{\Phi}_\tau^* d\tau$ and Ψ_t the fundamental solution of the Hamiltonian system given by:*

$$H(p,q) = \langle p, A_t q\rangle + \frac{1}{2}\left(\langle B_t B_t^* p, p\rangle + \langle R_t q, q\rangle\right)$$

Let $1+K$ be the second variation at said critical point, then:

$$\det(1+K) = (-1)^n e^{-n-2\int_0^t \operatorname{tr}(A_\tau) d\tau} \det(G^{-1}) \det(\Gamma^{-1}) \det(\Psi_t - \underline{\Phi}_t).$$

Proof. It is enough to complete the proof to evaluate $\mathfrak{p}_Q(s)$ at $s=1$ and the determinant of $\hat{\Phi}_t^* G_1 \hat{\Phi}_t$. For the first part we have:

$$\begin{aligned}\det(Q^1) &= \det(T_1^* J_1 \tilde{\Phi}_t^* \Phi_1^1 - T_0^* J_0) = \det(\underline{\Phi}_t^{-1} (\tilde{\Phi}_t)_* \Phi_1^1 - 1) \\ &= \det(\Psi_t - \underline{\Phi}_t).\end{aligned}$$

For the second claim notice that $\det(\hat{\Phi}_t^* G_1 \hat{\Phi}_t) = \det(G_1)\det(\hat{\Phi}_t)^2$ and $\det(\hat{\Phi}_t)$ solves the ODE

$$\frac{d}{dt}\det(\hat{\Phi}_t) = \det(\hat{\Phi}_t)\operatorname{tr}(A_t).$$

□

5.3 Proof of Hill's formula for general boundary conditions

In this section we provide a proof of Theorem 1.2. At first we work with separated boundary conditions and then reduce the general case to the former. The proof is a bit long so we try to give here a concise outline. The idea is to construct an analytic function f which vanishes precisely on the set $\{-1/\lambda : \lambda \in \operatorname{Spec}(K)\} \subseteq \mathbb{R}$. Particular care is needed to show that the multiplicity of the zeros of this function equals the multiplicity of the eigenvalues of K. We do this in Proposition 5.1 and Proposition 5.2 respectively. We show that this function decays exponentially and use a classical factorization Theorem by Hadamard to represent it as

$$f(s) = a s^k e^{bs} \prod_{\lambda \in \operatorname{Spec}(K)} (1+\lambda s)^{m(\lambda)}, \quad a, b \in \mathbb{C}, a \neq 0, k \in \mathbb{N}.$$

To prove the general case, we double the variables and consider general boundary conditions as separated ones. In this framework we compute the value of the parameters a, b and k appearing in the factorization.

5.3.1 Separated boundary conditions

We briefly recall the notation, we are working with an extremal λ_t with initial and final point $(\lambda_0, \lambda_1) \in Ann(N)$, where $N = N_0 \times N_1$ are the separated boundary conditions. We are assuming that λ_t is strictly normal and satisfies Legendre strong conditions. We work in a fixed tangent space, namely $T_{\lambda_0} T^*M$, to do so we backtrack our curve to its starting point λ_0 using the flow generated by the time dependent Hamiltonian:

$$h_{\tilde{u}}^t(\lambda) = \langle \lambda, f_{\tilde{u}(t)}(q)\rangle - \varphi_t(q, \tilde{u}(t)).$$

We denote the differential of said flow by $\tilde{\Phi}_*$. We have a scalar product g_i on $T_{\lambda_i}T^*M$, for each $i = 0, 1$. We assume that the orthogonal complement to the fibre at λ_i, $\Pi_i = T_{\lambda_i}T^*_{\pi(\lambda_i)}M$ is a Lagrangian subspace and that the range of Z_0 (and $\tilde{\Phi}_* Z_1$ respectively) is contained in $\Pi_0^\perp$ (resp. $\Pi_1^\perp$).

Remark 5.2. If we fix Darboux (i.e. *canonical*) coordinates coming from the splitting $\Pi_i \oplus \Pi_i^\perp$ it is straightforward to check that g_i takes a block diagonal form with symmetric $n \times n$ matrix G_i^j on the main diagonal. Similarly we can write down the coordinate representation of the matrix J_i and find:

$$g_i(X, Y) = \left\langle \begin{pmatrix} G_i^1 & 0 \\ 0 & G_i^2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \right\rangle, \quad J_i = \begin{pmatrix} 0 & -(G_i^1)^{-1} \\ (G_i^2)^{-1} & 0 \end{pmatrix}.$$

For $s \in \mathbb{R}$ (or $\mathbb{C}$) we introduce the following symplectic maps:

$$\begin{aligned} A_0^s(\eta) &= \eta + (1-s) J_0^{-1} pr_{\Pi_0^\perp} \eta, \quad \eta \in T_{\lambda_0}(T^*M), \\ A_1^s(\eta) &= \eta + (1-s)(J_1^{-1} + \tilde{\Phi}_* \circ pr_{\Pi_0} \circ \tilde{\Phi}_*^{-1}) pr_{\Pi_1^\perp} \eta, \quad \eta \in T_{\lambda_1}(T^*M). \end{aligned} \tag{5.7}$$

Notice that the transformation A_i^s are indeed symplectomorphisms. In (canonical) coordinates given by Π_i and $\Pi_i^\perp$ they have the following matrix representation:

$$A_i^s = \begin{pmatrix} 1 & (1-s) S_i \\ 0 & 1 \end{pmatrix}, \quad S_i^* = S_i.$$

The last map we are going introduce is two families of dilation in $T_{\lambda_0}(T^*M)$, one of the vertical subspace and one of its orthogonal complement. Let $s \in \mathbb{R}$ (or $\mathbb{C}$) and let us define the following maps:

$$\begin{aligned} \delta^s &: T_{\lambda_0}T^*M \to T_{\lambda_0}T^*M, \quad \delta^s \nu = s\, pr_{\Pi_0} \nu + pr_{\Pi_0^\perp} \nu \\ \delta_s &: T_{\lambda_0}T^*M \to T_{\lambda_0}T^*M, \quad \delta_s \nu = pr_{\Pi_0} \nu + s\, pr_{\Pi_0^\perp} \nu \end{aligned} \tag{5.8}$$

Proposition 5.1. *Let A_i^s be the maps given in eq. (5.7) and let Φ_1^s be the fundamental solution of the system:*

$$\dot{\eta} = Z_t^s (Z_t^s)^* J \eta, \quad Z_t^s = \delta^s Z_t,$$

The operator $1 + sK$ restricted to $\mathcal{V}$ has non trivial kernel if and only if there exists $(\eta_0, \eta_1) \in T_{(\lambda_0, \lambda_1)}(Ann(N))$ such that

$$A_1^s \circ \tilde{\Phi}_* \circ \Phi_1^s \circ A_0^s \eta_0 = \eta_1$$

In particular, the geometric multiplicity of the kernel of $1 + sK$ equals the number of linearly independent solutions of the above equation.

Proof. The equation for the kernel becomes $\langle u, v \rangle + \langle sKu, v \rangle = 0$ $\forall u, v \in \mathcal{V}$ which means $u + sKu \in \mathcal{V}^\perp$ i.e. (see Lemma 5.1 for a description of $\mathcal{V}^\perp$):

$$\begin{cases} u_0 = s u_0 + v_0 \Rightarrow (1-s) u_0 = v_0 \\ u_t = s Z_t^* J \Big(\int_0^t Z_\tau u_\tau d\tau + Z_0 u_0 \Big) + Z_t^* J \nu \\ (1-s) u_1 = s \Lambda(u) + v_1 \end{cases} \tag{5.9}$$

Let us substitute Z_t with $Z_t^s = \delta^s Z_t$. It is immediate to check, using the definition in eq. (5.8), that:

$$sZ_t^* J \int_0^t Z_\tau u_\tau d\tau = (Z_t^s)^* J \int_0^t Z_\tau^s u_\tau d\tau.$$

Moreover $Z_t^* J\nu = (Z_t^s)^* J\nu$ and $sZ_t^* JZ_0 = (Z_t^s)^* JZ_0^s$ since $\delta^s Z_0 = Z_0$.

All the calculation we will do from here on are aimed at rewriting eq. (5.9) as a boundary value problem in $T_{\lambda_0}T^*M \times T_{\lambda_0}T^*M$. Let us start with the second equality in eq. (5.9) and set

$$\eta(t) = \int_0^t Z_\tau^s u_\tau d\tau + Z_0^s u_0 + \nu, \quad \eta(0) = Z_0^s u_0 + \nu.$$

The linear constraint defining $\mathcal{V}$ implies:

$$\begin{aligned} Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1 \in \Pi &\iff Z_0 u_0 + \int_0^1 Z_t^s u_t dt + Z_1^s u_1 \in \Pi \\ &\iff \eta(1) + Z_1^s u_1 \in \Pi_0 \\ &\iff \tilde{\Phi}_*(\eta(1) + Z_1^s u_1) \in \Pi_1 \end{aligned}$$

This means we are looking for solution that have initial condition in $\pi_*^{-1}TN_0$ and final value in $\pi_*^{-1}TN_1$. If we multiply by Z_t^s the second equation in eq. (5.9), we are brought to consider the following problem:

$$\begin{cases} \dot{\eta}(t) = Z_t^s (Z_t^s)^* J\eta(t) \\ (\pi_*\eta(0), \pi_*\eta(1)) \in T(N_0 \times N_1) \end{cases} \tag{5.10}$$

Now we use the remaining equations in eq. (5.9) to reduce the space $\pi_*^{-1}(T(N_0 \times N_1))$ to a Lagrangian one. Let us regard the first and third line in eq. (5.9) as equations in $T_{\lambda_0}T^*M$ and $T_{\lambda_1}T^*M$. Using the maps Z_0 and $\tilde{\Phi}_* Z_1$ we obtain:

$$\begin{cases} (1-s)Z_0 u_0 = Z_0 v_0 = pr_0 J_0 \nu \\ (1-s)\tilde{\Phi}_* Z_1 u_1 = \tilde{\Phi}_* Z_1 v_1 + s\tilde{\Phi}_* Z_1 \Lambda(u) = pr_1 J_1 \tilde{\Phi}_* \nu + s\tilde{\Phi}_* Z_1 \Lambda(u) \end{cases}$$

Notice that, for $u \in \mathcal{V}$ we have that:

$$s(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1) = Z_0 u_0 + \int_0^1 Z_t^s u_t dt + Z_1^s u_1$$

This implies that the term $s\tilde{\Phi}_* Z_1 \Lambda(u)$ can be rewritten as:

$$s\tilde{\Phi}_* Z_1 \Lambda(u) = s\, pr_1 J_1 \tilde{\Phi}_*(Z_0 u_0 + \int_0^1 Z_t u_t dt + Z_1 u_1) = pr_1 J_1 \tilde{\Phi}_*(\eta(1) + Z_1^s u_1 - \nu)$$

If substitute $Z_0 u_0$ with $pr_{\Pi_0^\perp}\eta(0)$ we end up with the equations:

$$\begin{cases} (1-s)pr_{\Pi_0^\perp}\eta(0) = pr_0 J_0 \nu = pr_0 J_0 \eta(0) \\ (1-s)\tilde{\Phi}_* Z_1 u_1 = pr_1 J_1 \tilde{\Phi}_*(\eta(1) + Z_1^s u_1) \end{cases}$$

Now we do the same kind of substitution for the term $Z_1^s u_1$. Using the projections on Π_0 and $\Pi_0^\perp$ and recalling that $\tilde{\Phi}_*$ sends Π_0 to Π_1, we have:

$$\begin{aligned} Z_1^s u_1 &= s\, pr_{\Pi_0} Z_1 u_1 + pr_{\Pi_0^\perp} Z_1 u_1 = (s-1)\, pr_{\Pi_0} Z_1 u_1 + Z_1 u_1 \\ pr_1 J_1 \tilde{\Phi}_* Z_1^s u_1 &= (s-1)\, pr_1 J_1 \tilde{\Phi}_* pr_{\Pi_0} Z_1 u_1 + pr_1 J_1 \tilde{\Phi}_* Z_1 u_1 \\ &= (s-1)\, pr_1 J_1 \tilde{\Phi}_* pr_{\Pi_0} Z_1 u_1 \end{aligned}$$

Last equality being due to the fact that the image of $\tilde{\Phi}_* Z_1$ is isotropic and thus $J_1 \operatorname{Im}(\tilde{\Phi}_* Z_1) \subset \operatorname{Im}(\tilde{\Phi}_* Z_1)^\perp$. Moreover $\tilde{\Phi}_* Z_1 u_1$ coincides with the projection of $-\tilde{\Phi}_* \eta(1)$ on $\Pi_1^\perp$. Thus we are left with equations:

$$\begin{cases} (1-s) pr_{\Pi_0^\perp} \eta(0) = pr_0 J_0 \nu = pr_0 J_0 \eta(0) \\ (1-s)(-pr_{\Pi_1^\perp} \tilde{\Phi}_* \eta(1) + pr_1 J_1 \tilde{\Phi}_* pr_{\Pi_0} Z_1 u_1) = pr_1 J_1 \tilde{\Phi}_* \eta(1) \end{cases} \tag{5.11}$$

It is straightforward to check that $pr_1 J_1 \tilde{\Phi}_* pr_{\Pi_0} Z_1 u_1$ depends only on the projection of $Z_1 u_1$ on $\Pi_0^\perp$. Moreover expanding $1 = \tilde{\Phi}_* \circ \tilde{\Phi}_*^{-1}$ and using the relation $\tilde{\Phi}_* Z_1 u_1 = -pr_{\Pi_1^\perp} \tilde{\Phi}_* \eta(1)$, the second equality in eq. (5.11) can be rewritten as:

$$(s-1) pr_{\Pi_1^\perp} \tilde{\Phi}_* \eta(1) = pr_1 J_1 \tilde{\Phi}_* \eta(1) + (1-s) pr_1 J_1 \tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_*^{-1} pr_{\Pi_1^\perp} \tilde{\Phi}_* \eta(1) \tag{5.12}$$

If $s = 1$ we have that $pr_0(J_0 \eta(0))$ and $pr_1 J_1 \tilde{\Phi}_* \eta(1)$ are zero. Consider the first case, the equation is equivalent to:

$$\sigma(\eta(0), Z_0 w_0) = g_0(J_0 \eta(0), Z_0 w_0) = g_0(pr_0 J_0 \eta(0), Z_0 w_0) = 0,\ \forall w_0 \in \mathbb{R}^{\dim(N_0)}.$$

Thus we are looking for solution starting from $T_{\lambda_0} Ann(N_0)$. Similarly setting $s = 1$ in the second equality we find that $\eta(1)$ must lie inside $T_{\lambda_1} Ann(N_1)$.

Now, we want to interpret the boundary conditions as an analytic family of Lagrangian subspaces depending on s. To do so we employ the following linear map defined in eq. (5.7):

$$A_0^s(\eta) = \eta + J_0^{-1}(1-s) pr_{\Pi_0^\perp} \eta$$

If $\eta \in T_{\lambda_0} Ann(N_0)$ we have that $pr_{\Pi_0^\perp} \eta = pr_0 \eta$ and $pr_0 J_0 \eta = 0$ and thus:

$$\begin{aligned} pr_{\Pi_0^\perp}(A_0^s(\eta)) &= pr_{\Pi_0^\perp}(\eta) \in \operatorname{Im}(Z_0) \\ pr_0 J_0 (A_0^s(\eta)) &= pr_0(J_0 \eta + (1-s) pr_{\Pi_0^\perp} \eta) \\ &= pr_0 J_0 \eta + (1-s) pr_{\Pi_0^\perp} \eta \\ &= (1-s) pr_{\Pi_0^\perp}(A_0^s(\eta)) \end{aligned}$$

So we have shown that $A_0^s(T_{\lambda_0} Ann(N_0))$ is precisely the space satisfying the first set of equations. A similar argument works for the final point. Let us recall the definition of A_1^s given in eq. (5.7):

$$A_1^s(\eta) = \eta + (1-s)(J_1^{-1} + \tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_*^{-1}) pr_{\Pi_1^\perp} \eta.$$

Now we check that the boundary condition for the final point are satisfied if $A_1^s \circ \tilde{\Phi}_*\, \eta(1) \in T_{\lambda_1} Ann(N_1)$. In fact, take any η in $T_{\lambda_1} Ann(N_1)$, it holds:

$$\begin{aligned} pr_{\Pi_1^\perp} (A_1^s)^{-1} \eta &= pr_{\Pi_1^\perp} \eta, \\ pr_{\Pi_1} (A_1^s)^{-1} \eta &= pr_{\Pi_1} \eta + (s-1)(J_1^{-1} + \tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_*^{-1}) pr_{\Pi_1^\perp} \eta, \\ pr_1 J_1 (A_1^s)^{-1} \eta &= (s-1) pr_1 (1 + J_1 \tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_*^{-1}) pr_{\Pi_1^\perp} \eta. \end{aligned}$$

It is now straightforward to substitute the last equality in eq. (5.12) and check that indeed $(A_1^s)^{-1}(T_{\lambda_1}Ann(N_1))$ is the right space.

Let us call Φ_1^s the fundamental solution of eq. (5.10) at time 1 and denote by $\Gamma(\Phi_1^s)$ its graph. It follows that $s \in \mathbb{R} \setminus \{0\}$ is in the kernel of $1 + sK$ if and only if:

$$\Gamma(\tilde{\Phi}_* \circ \Phi_1^s) \cap A_0^s(T_{\lambda_0}Ann(N_0)) \times (A_1^s)^{-1}(T_{\lambda_1}Ann(N_1)) \neq (0)$$

which is equivalent to the condition:

$$\Gamma(A_1^s \circ \tilde{\Phi}_* \circ \Phi_1^s \circ A_0^s) \cap T_{\lambda_0}Ann(N_0) \times T_{\lambda_1}Ann(N_1) \neq (0) \tag{5.13}$$

Now we prove the part about the multiplicity. Suppose that two different controls u and v give the same trajectory η_t solving eq. (5.10). Since the maps Z_0 and Z_1 are injective it must hold that $v_0 = u_0$ and $v_1 = u_1$. Moreover $\int_0^t Z_\tau u_\tau d\tau = \int_0^t Z_\tau v_\tau d\tau$ and thus:

$$Ku = Z_t^* J \int_0^t Z_\tau u_\tau d\tau = Z_t^* J \int_0^t Z_\tau v_\tau d\tau = Kv.$$

However Volterra operator are always injective and thus $u = v$.

Vice-versa consider $u = 0$ and see whether you get solution of the system above that do not correspond to any variation. Since u_0 and u_1 are both zero we are considering solution starting from the fibre and reaching the fibre. Plugging in $u_t = 0$ we obtain:

$$0 = \dot{\eta} = Z_t^s (Z_t^s)^* J\eta = Z_t^s (Z_t^s)^* J\nu$$

However $pr_{\Pi^\perp} Z_t^s (Z_t^s) J\nu = X_t X_t^* \nu$ and by assumption the matrix $\int_0^1 X_t X_t^* dt$ is invertible. Thus we get a contradiction. □

Remark 5.3. If we complexify all the subspaces involved in the proof of Proposition 5.1, i.e. tensor with $\mathbb{C}$ we can take also $s \in \mathbb{C}$.

We can reformulate the intersection problem in the statement of Proposition 5.1 as follows. Let $\tilde{\pi}$ the orthogonal projection, with respect to g_1, onto the subspace $T_{\lambda_1}Ann(N_1)^\perp$ and define a map Q^s as:

$$Q^s : T_{\lambda_0}Ann(N_0) \to T_{\lambda_1}Ann(N_1)^\perp, \quad Q^s(\eta) = \tilde{\pi}(A_1^s \Phi_1^s A_0^s)(\eta). \tag{5.14}$$

Let us fix now two bases, one of $T_{\lambda_0}Ann(N_0)$ and one of $T_{\lambda_1}Ann(N_1)$. Construct two $2n \times n$ matrices using the elements of the chosen basis, let us call the resulting objects T_0 and T_1 respectively. It follows that $J_1 T_1$ is a base of $T_{\lambda_1}Ann(N_1)^\perp$. Define the scalar function $\det(Q^s)$ as the determinant of the $n \times n$ matrix $T_1^* J_1 A_1^s \Phi_1^s A_0^s T_0$. Clearly different choices of basis give simply a scalar multiple of $\det(Q^s)$ and thus is well defined:

$$\mathfrak{p}_Q(s) = \det(Q^s) = \frac{\det(T_1^* J_1 A_1^s \Phi_1^s A_0^s T_0)}{\det(T_0^* T_0)^{1/2} \det(T_1^* J_1 J_1^* T_1)^{1/2}}$$

Moreover $\det(Q^s)|_{s=s_0} = 0$ if and only if there exists at least a solution to our boundary problem. Notice that map $s \mapsto \det(Q^s)$ is analytic in s since the fundamental matrix is an entire map in s (see [7][Proposition 4]). The following Proposition shows that the multiplicity of any root $s_0 \neq 0$ is equal to the number of independent solutions to the boundary value problem.

Proposition 5.2. *The multiplicity of any root $s_0 \neq 0$ of $\det Q^s$ equals the dimension of the kernel of Q^s.*

Proof. The proof is done in two steps. First of all we show that the equation $\det(Q^s) = 0$ is equivalent to $\det(R^s) = 0$ where R^s is a symmetric matrix, analytic in s. Once one knows this, it suffice to compute $\partial_s R^s$ and show that it is non degenerate to show that the multiplicity of the equation is the same as the dimension of the kernel.

Step 1: Replace Q^s with a symmetric matrix

First of all we can consider $s \in \mathbb{R}$ since all the roots are real. As remarked above the determinant of the matrix Q^s is zero whenever the graph of $A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s$ intersect the subspace $L_0 = T_{(\lambda_0,\lambda_1)}(Ann(N_0 \times N_1))$. Suppose that s_0 is a time of intersection and choose as coordinates in the Lagrange Grassmannian L_0 and another subspace L_1 transversal to both L_0 and $\Lambda_s := \Gamma(A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s)$. This means that, if $(T_{\lambda_0}T^*M)^2 \approx \{(p,q) | p, q \in \mathbb{R}^{2n}\}$ we can identify $L_0 = \{q = 0\}$ and $L_1 = \{p = 0\}$.

In this coordinates Λ_s is given by the graph of a symmetric matrix, i.e. is the following subspace $\Lambda_s = \{(p, R^s p)\}$ where again R^s is analytic in s.

The quadratic form associated to the derivative $\partial_s R(s)$ can be interpreted as the velocity of the curve $s \mapsto \Lambda_s$ inside the Grassmannian, it is possible to compute it choosing an arbitrary base of Λ_s and an arbitrary set of coordinates. Invariants such as signature and nullity do not change (see for example [10, 14] or [6]). Take a curve $\lambda_s = (p_s, R^s p_s)$ inside Λ_s then one has:

$$S(\lambda_s) = \sigma(\lambda_s, \dot{\lambda}_s) = \langle p_s, \partial_s R^s p_s \rangle$$

Recall that we will be using the symplectic form given by $(-\sigma_{\lambda_0}) \oplus \sigma_{\lambda_1}$, in order to have that graph of a symplectic map is a Lagrangian subspace.

Step 2: Replace Λ_s with a positive curve

We slightly modify our curve to exploit an hidden positivity of the Jacobi equation. We substitute the fundamental solution Φ_1^s with the following map:

$$\Psi^s = \Psi_1^s = \delta_s \Phi_1^s \delta_{\frac{1}{s}}.$$

It is straightforward to check that Ψ^s is again a symplectomorphism and that it is the fundamental solution of the following ODE system at time $t = 1$:

$$\dot{\Psi}_t^s = s Z_t Z_t^* J \Psi_t^s, \quad \Psi_0^s = Id. \tag{5.15}$$

On one hand we are introducing a singularity at $s = 0$ but on the other hand we are going to show that the graph of Ψ^s becomes a monotone curve and its velocity is fairly easy to compute.

First of all, hoping that the slight abuse of notation does not create any confusion, let us introduce a family of dilations similar to the δ_s,δ^s also in $T_{\lambda_1}T^*M$. The definition is analogous to the one in eq. (5.8) but with Π_1 and $\Pi_1^\perp$ and we will denote them with the same symbol.

Let us consider the following symplectomorphisms:

$$\delta_s A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s \delta_{\frac{1}{s}} = \delta_s A_1^s \tilde{\Phi}_* \delta_{\frac{1}{s}} \Psi_s \delta_s A_0^s \delta_{\frac{1}{s}}$$

Notice that the dilations δ_s preserve the subspaces $T_{\lambda_i}Ann(N_i)$ and thus the intersection points between the graph of the above map and the subspace $T_{(\lambda_0,\lambda_1)}Ann(N_0 \times N_1)$ are unchanged. Let us rewrite the maps $\delta_s A_0^s \delta_{\frac{1}{s}}$ and $\delta_s A_1^s \tilde{\Phi}_* \delta_{\frac{1}{s}}$. For the former:

$$\delta_s A_0^s \delta_{\frac{1}{s}} = \delta_s(1 + (1-s)J_0^{-1} pr_{\Pi_0^\perp})\delta_{\frac{1}{s}} = 1 + \frac{1-s}{s} J_0^{-1} pr_{\Pi_0^\perp} = B_0^s$$

For the latter a computation in local coordinates and the fact that the dilations δ_s and $\tilde{\Phi}_*$ do not commute yield:

$$\begin{aligned} \delta_s A_1^s \tilde{\Phi}_* \delta_{\frac{1}{s}} &= \delta_s(1 + (1-s)(J_1^{-1} + \tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_*^{-1}) pr_{\Pi_1^\perp}) \tilde{\Phi}_* \delta_{\frac{1}{s}} \\ &= (1 + \frac{1-s}{s} J_1^{-1} pr_{\Pi_1^\perp}) \tilde{\Phi}_* = B_1^s \tilde{\Phi}_* \end{aligned}$$

Thus we take, for $s \neq 0$, as curve $\Lambda_s := \Gamma(B_1^s \tilde{\Phi}_* \Psi^s B_0^s)$, the graph of the symplectomorphism just introduced. Notice that Ψ^s is actually analytic, the singularity at $s = 0$ come only form the maps B_i^s.

Step 3: Computation of the velocity

Now we compute the velocity of the graph of $B_1^s \tilde{\Phi}_* \Psi^s B_0^s$. Take a curve

$$\lambda_s = (\eta, B_1^s \tilde{\Phi}_* \Psi^s B_0^s \eta)$$

inside Λ_s and let us compute the quadratic form associated to the velocity:

$$\begin{aligned} S(\lambda_s) &= -\sigma(\eta, \partial_s \eta) + \sigma(B_1^s \tilde{\Phi}_* \Psi^s B_0^s \eta, \partial_s(B_1^s \tilde{\Phi}_* \Psi^s B_0^s \eta)) \\ &= \sigma(B_1^s \tilde{\Phi}_* \Psi^s B_0^s \eta, (\partial_s B_1^s) \tilde{\Phi}_* \Psi^s B_0^s \eta) + \sigma(\Psi^s B_0^s \eta, (\partial_s \Psi^s) B_0^s \eta) + \\ &\quad + \sigma(B_0^s \eta, \partial_s(B_0^s)\eta) \end{aligned}$$

Let us consider the terms of the type $\sigma(B_i^s x, \partial_s B_i^s x)$. It immediate to compute the derivative in this case, recall that $B_i^s x = x + \frac{(1-s)}{s} J_i^{-1} pr_{\Pi_i^\perp} x$. It follows that $\partial_s B_i^s = -\frac{1}{s^2} J_i^{-1} pr_{\Pi_i^\perp}$ thus the first and last term read as:

$$\begin{aligned} \sigma(B_1^s \xi, (\partial_s B_1^s)\xi) &= -\frac{1}{s^2}\sigma(\xi, J_1^{-1} pr_{\Pi_1^\perp} \xi) \\ &= \frac{1}{s^2} g_1(pr_{\Pi_1^\perp}\xi, pr_{\Pi_1^\perp}\xi), \quad \text{where } \xi = \tilde{\Phi}_* \Psi^s B_0^s \eta, \\ \sigma(\partial_s(B_0^s)\eta, B_0^s \eta) &= -\frac{1}{s^2}\sigma(B_0^s \eta, J_0^{-1} pr_{\Pi_0^\perp}\eta) = -\frac{1}{s^2} g_0(J_0 \eta, J_0^{-1} pr_{\Pi_0^\perp}\eta) \\ &= \frac{1}{s^2} g_0(pr_{\Pi_0^\perp}\eta, pr_{\Pi_0^\perp}\eta). \end{aligned}$$

Notice we used the fact that J_i (and thus J_i^{-1}) is g_i-skew symmetric. Now we rewrite the middle term. We present it as the integral of its derivative using the equation for Ψ_t^s. Let us use the shorthand notation $x = B_0^s \eta$. We obtain:

$$\begin{aligned} \frac{d}{dt}\left(\sigma(\Psi_t^s x, (\partial_s \Psi_t^s)x)\right) &= \sigma(\partial_t \Psi_t^s x, (\partial_s \Psi_t^s)x) + \sigma(\Psi_t^s x, (\partial_s \partial_t \Psi_t^s)x) \\ &= s\sigma(Z_t Z_t^* J \Psi_t^s x, \partial_s \Psi_t^s x) + s\sigma(\Psi_t^s x, Z_t Z_t^* J \partial_s \Psi_t^s x) \\ &\quad + \sigma(\Psi_t^s x, \partial_s\big(s Z_t Z_t^*\big) J \Psi_t^s x). \end{aligned}$$

The first and second term have opposite sign and thus cancel out. What remains is:

$$\frac{d}{dt}\left(\sigma((\partial_s \Psi_t^s)x, \Psi_t^s x)\right) = \sigma(\Psi_t^s x, Z_t Z_t^* J \Psi_t^s x) = g(Z_t^* J \Psi_t^s x, Z_t^* J \Psi_t^s x)$$

Integrating over $[0,1]$ and using the fact that $\partial_s \Psi_t^s|_{(0,0)} = 0$ we get that:

$$\sigma(\partial_s \Psi^s x, \Psi^s x) = \int_0^1 g(Z_t^* J \Psi_t^s x, Z_t^* J \Psi_t^s x)dt.$$

Using the notation $||\cdot||$ to denote the norm with respect to the corresponding metric and summing everything up we find the following expression for the velocity of our curve:

$$S_s(\lambda_s) = \frac{1}{s^2}\left(\|pr_{\Pi_0^\perp}\eta\|^2 + \|pr_{\Pi_1^\perp}\Psi^s B_0^s \eta\|^2\right) + \int_0^1 \|Z_t^* J \Psi_t^s B_0^s \eta\|^2 dt$$

Since each term of the sum is non positive $S_s(\lambda_s)$ is zero if and only if each term is zero. From the first one we obtain that η must be contained in the fibre. Notice that B_0^s acts as the identity on Π_0 and thus in this case $\|pr_{\Pi_1^\perp}\tilde{\Phi}_* \Psi^s B_0^s \eta\|^2 = \|pr_{\Pi_1^\perp}\Psi^s \eta\|^2$.

It follows that $\Psi_t^s B^s \eta = \Psi_t^s \eta$ is a solution of the Jacobi equation (5.10) starting and reaching the fibre (recall that $\tilde{\Phi}_*(\Pi_0) = \Pi_1$). Let us consider now the third piece, since the integrand is positive it must hold that for almost any t, $Z_t^* J \Psi_t^s \eta = 0$. If we multiply this equation by Z_t we find that:

$$Z_t Z_t^* J \Psi_t^s \eta = 0 = \dot{\Psi}_t^s \eta.$$

It follow that we are dealing with a constant solution starting and reaching the fibre. However this contradicts the assumption that the matrix $\int_0^1 X_t X_t^* dt$ is non degenerate. In fact, if we substitute a non zero constant solutions starting from the fibre in eq. (5.10), we find that $pr_{\Pi_0^\perp}(\eta) = \int_0^1 X_t X_t^* dt\, \eta \neq 0$ $\square$

The following proposition is proved in [7].

Proposition 5.3. *There exists $c_1, c_2 > 0$ such that:*

$$\|\Phi_1^s\| \le c_1 e^{c_2|s|} \quad \forall\, s \in \mathbb{C}$$

Moreover Φ_t^s is analytic and the function $s \mapsto \det(Q^s)$ is entire and satisfy the same type of estimate.

This fact tell us that $\det Q^s$ is an entire function of order $\rho \le 1$. We know its zeros which are determined by the eigenvalues of K and thus we can apply Hadamard factorization theorem ([26]) to present it as an infinite product. It follows that we have the following identity:

$$\det(Q^s) = a s^k e^{bs} \prod_{\lambda \in Sp(K)} (1 + s\lambda)^{m(\lambda)} \quad a, b \in \mathbb{C}, a \neq 0, k \in \mathbb{N} \tag{5.16}$$

where $m(\lambda)$ is the geometric multiplicity of the eigenvalue λ. To determine the remaining parameters it sufficient to know the value of $\det(Q^s)$ and a certain number of its derivatives at $s = 0$ (depending on the value of k). Assume for now that $k = 0$, a straightforward computation shows that:

$$\det(Q^s)|_{s=0} = a, \quad \partial_s \det(Q^s)|_{s=0} = a(b + \mathrm{tr}(K)).$$

We will compute these quantities in the next section for a general set of boundary condition manifold N.

5.3.2 General boundary condition

In this section we use Proposition 5.1 to prove a determinant formula for general boundary conditions $N \subseteq M \times M$. We reduce this case to the case of separate boundary conditions.

Let us consider $M \times M$ as state space, with the following dynamical system:

$$f_u(q', q) = \begin{pmatrix} 0 \\ f_u(q) \end{pmatrix}, \ (q', q) \in M \times M.$$

and boundary conditions $\Delta \times N$. With this definition, any extremal between two points q_0 and q_1 lifts naturally to an extremal between (q_0, q_0) and (q_0, q_1). However the end point map of the new system is no longer surjective since any trajectory is confined to a submanifold of the form $\{\hat{q}\} \times M$. We have to slightly modify the arguments of the previous section but everything remains essentially unchanged.

First of all notice that Pontryagin maximum principle implies that the lift of the extremal curve $\tilde{q}(t) = (q_0, q(t))$ is the curve $\tilde{\lambda}(t) = (-\lambda_0, \lambda(t))$. This is because the initial and final covector of the lift must annihilate the tangent space of the boundary conditions manifold. In this case $N_0 = \Delta$ and the annihilator of the diagonal subspace is $\{(\lambda, -\lambda) : \lambda \in T_{\lambda_0} T^*_{q_0} M\}$. Moreover, by the orthogonality condition in PMP (see Theorem B.1), we know that $(-\lambda(0), \lambda(1))$, the initial and final points of the original extremal, annihilate the tangent space of N.

Thus, if we want to work in one fixed tangent space, we have to multiply the first covector by -1. This changes the sign of the symplectic form and we are thus brought to work on $T_{\lambda_0} T^*M \times T_{\lambda_0} T^*M$ with symplectic form $(-\sigma) \oplus \sigma$.

With this change of sign, the tangent space to the annihilator of the diagonal gets mapped to the diagonal subspace of $T_{\lambda_0} T^*M \times T_{\lambda_0} T^*M$. The tangent space to the annihilator of the boundary conditions N is mapped to the tangent space of:

$$A(N) = \{(\mu_0, \mu_1) : \langle \mu_0, X_0 \rangle - \langle \mu_1, X_1 \rangle = 0, \ \forall (X_0, X_1) \in TN\}.$$

Let A_i^s be the map given in eq. (5.7). The following proposition is the counter part of Proposition 5.1 for general boundary conditions.

Proposition 5.4. *Let Φ_1^s be the fundamental solution of the Jacobi system:*

$$\dot{\eta} = Z_t^s (Z_t^s)^* J \eta.$$

The operator $1 + sK$ restricted to $\mathcal{V}$ has non trivial kernel if and only if there exists $(\eta_0, \eta_1) \in T_{(\lambda_0, \lambda_1)} A(N)$ such that

$$A_1^s \circ \tilde{\Phi}_* \circ \Phi_1^s \circ A_0^s \eta_0 = \eta_1 \tag{5.17}$$

The geometric multiplicity of the kernel equals the number of linearly independent solutions of the above equation.

Proof. We apply Proposition 5.1 to the new problem. As already mentioned, since we are working with the symplectic form $(-\sigma) \oplus \sigma$, we have to substitute the annihilator of the diagonal with:

$$A(\Delta) = \{(\lambda, \lambda) : \lambda \in T^*M\} \Rightarrow T_{\lambda_0}(A(\Delta)) = \{(\eta, \eta) : \eta \in T_{\lambda_0} T^*M\}$$

We have also more freedom in choosing the Riemannian metrics g_0 and g_1 at the boundary points. Since we are working with a product space it seems reasonable to take the sum of metrics defined on each factor. Set $g_{0,j}$ to be the metric on the $j-$th component of the initial point and $g_{1,j}$ the metric on the $j-$th component of the final one.

Define the maps $\hat{A}_i^s$ in the same fashion as in eq. (5.7) but on the bigger space $(T_{\lambda_i}T^*M)^2$. Since we chose $g_i = g_{i,1} \oplus g_{i,2}$ these maps split as the direct sums of two symplectomorphisms. We have $\hat{A}_i^s = (A_{i,1}^s)^{-1} \oplus A_{i,2}^s$ for $i = 0,1$. Writing this in matrix form means:

$$\hat{A}_i^s = \begin{pmatrix} (A_{i,1}^s)^{-1} & 0 \\ 0 & A_{i,2}^s \end{pmatrix}$$

Notice that we have to take the inverse on the first component to be consistent with the definition in eq. (5.7) because of the different sign of the symplectic form.

The solution to the Jacobi system is the same as the original one on the second component and the identity on the first one. Thus the second variation has a kernel if and only if:

$$\hat{A}_1^s \begin{pmatrix} 1 & 0 \\ 0 & \tilde{\Phi}_* \Phi_1^s \end{pmatrix} \hat{A}_0^s \begin{pmatrix} \eta \\ \eta \end{pmatrix} \in T_{(\lambda_0,\lambda_1)}A(N), \quad \eta \in T_{\lambda_0}T^*M$$

It is straightforward to get rid off the diagonal and write the equations in terms of the intersection of a graph and of the annihilator of the boundary conditions. Take $(\eta_0, \eta_1) \in T_{(\lambda_0,\lambda_1)}A(\Delta)$, then $(A_{0,1}^s)(A_{1,1}^s)\eta_0 = \eta$. Inserting in the equation above we have that $A_{1,2}^s \tilde{\Phi}_* \Phi_1^s A_{0,2}^s A_{0,1}^s A_{1,1}^s \eta_0 = \eta_1$.

Notice that in the case we are considering the map $A_{1,1}^s$ has the same structure as the maps $A_{0,j}^s$ since $\tilde{\Phi}_*$ acts as the identity on the first component and the extra piece $\tilde{\Phi}_* pr_{\Pi_0} \tilde{\Phi}_* pr_{\Pi_0^\perp}$ vanishes. In particular it depends only on the metric we are fixing on the final point. So it is clear that $A_{0,2}^s A_{0,1}^s A_{0,2}^s$ depends only on the sum of the corresponding metrics. Thus we recover precisely the formula in the statement.

Notice that, again, to trivial variations correspond constant solutions starting from the fibre and reaching the fibre. So in principle we would have to factor out the trivial dynamic on the first component. However the boundary condition $(\eta, \eta) \in \Pi \cap A(\Delta)$ implies that any such constant solution must have the same initial condition. Since there are no constant solution starting from the fibre for the original Jacobi system it follows that the equation $A_1^s \circ \tilde{\Phi}_* \circ \Phi_1^s \circ A_0^s \eta_0 = \eta_1$ has the same number of independent solutions as the dimension of kernel of $1 + sK$. □

Now we define an analogous map to the one in eq. (5.14). Let π_N be the orthogonal projection on the space $T_{(\lambda_0,\lambda_1)}A(N)^\perp$ and consider the map:

$$Q^s : \Gamma(A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s) \to T_{(\lambda_0,\lambda_1)}A(N)^\perp, \quad Q^s(\eta) = \pi_N(\eta). \tag{5.18}$$

Let $T = (T_0, T_1)$ be any linear invertible map from $\mathbb{R}^{2n}$ to the space $T_{(\lambda_0,\lambda_1)}A(N)$. We denote by J the map $(-J_0) \oplus J_1$ representing the symplectic form $(-\sigma_{\lambda_0}) \oplus \sigma_{\lambda_1}$. As in the previous section we define the following scalar function:

$$\det(Q^s) = \frac{\det(T_1^* J_1 A_1^s \tilde{\Phi}_* \Phi_1^s A_0^s - T_0^* J_0)}{\det(T_0^* J_0 J_0^* T_0 + T_1^* J_1 J_1^* T_1)^{1/2}}$$

Remark 5.4. The two definition given by eq. (5.18) and eq. (5.14) agree (up to a positive scalar) whenever the boundary conditions are of the type $N_0 \times N_1$. Moreover, eq. (5.18) is essentially the same definition given in eq. (5.14) applied to the extended system introduced at the beginning of this section, where we take as boundary conditions $\Delta \times N$ (obviously again, up to a positive scalar).

Proposition 5.5. *The multiplicity of any roots $s_0 \neq 0$ of the equation $\det(Q^s)$ is equal to the geometric multiplicity of the boundary value problem.*

Proof. First of all notice that the same proof of Proposition 5.2 works for the curve of the extended system verbatim. The only slight difference is that the velocity of the curve is non degenerate only on the intersection with $TA(N)$. But this is still sufficient to conclude. $\square$

In the remaining part of this section we carry out the computation of the normalizing factors of the function $\det(Q^s)$. As already mentioned at the end of the previous section a classical factorization theorem by Hadamard (see [26]) tells us that:

$$\det(Q^s) = as^k e^{bs} \prod_{\lambda \in Sp(K)} (1+s\lambda)^{m(\lambda)} \quad a, b \in \mathbb{C}, a \neq 0, k \in \mathbb{N}$$

where $m(\lambda)$ is the geometric multiplicity of the eigenvalue. We are now going to compute the values of $a, b \in \mathbb{C}$ and k.

Theorem 5.3. *For almost any choice of metrics g_0, g_1 on $T_{\lambda_i}T^*M$, $\det(Q^s|_{s=0}) \neq 0$. Whenever this condition holds, the determinant of the second variation is given by:*

$$\det(1+sK) = \det((Q^s|_{s=0})^{-1}) e^{s(\operatorname{tr}(K) - \operatorname{tr}(\partial_s Q^s\,(Q^s)^{-1}|_{s=0}))} \det(Q^s) \tag{5.19}$$

Proof. First of let us show that, for almost any choice of scalar product, $k = 0$ and thus $a = \det(Q_1^s|_{s=0}) \neq 0$. This is equivalent to a transversality condition between the graph of the symplectomorphism $A_1^s \tilde{\Phi}_* \Phi^s A_0^s$ and the annihilator of the boundary conditions N.

We can argue as follows: consider the following family of maps acting on the Lagrange Grassmannian of $T_{\lambda_0}T^*M \times T_{\lambda_1}T^*M$ depending on the choice of scalar products G_0 and G_1:

$$F_G = (A_0^s)^{-1} \times A_1^s|_{s=0}, \quad G = (G_0, G_1), G_i > 0.$$

It is straightforward to see that they define a family of algebraic maps of the Grassmannian to itself. For any fixed subspace L_0, we have that $F_G^{-1}(L_0)$ is arbitrary close to $\Pi_0 \times \Pi_1$, for G_i large enough. Notice that $\Gamma(A_1^s \tilde{\Phi}_* \Phi^s A_0^s) \cap L_0 \neq (0)$ if and only if $\Gamma(\tilde{\Phi}_* \Phi^s) \cap F_G^{-1}(L_0) \neq (0)$. It is straightforward to check using the formula in Lemma 5.4 that $\Gamma(\tilde{\Phi}_* \Phi^s)$ is transversal to $\Pi_0 \times \Pi_1$ and thus to $F_G^{-1}(L_0)$ for any fixed L_0 and G_i sufficiently large. Now, since everything is algebraic in G and there is a Zariski open set in which we have transversality, the possible choices of G_i for which $k > 0$ are in codimension 1.

Let us assume that $k = 0$ and compute b. Differentiating the expression for $\det(Q^s)$ in eq. (5.16) at $s = 0$ we find that:

$$\partial_s \det(Q^s)|_{s=0} = a(b + \operatorname{tr}(K))$$

A integral formula for the trace of K is given in Lemma 5.3. The derivative of $\det(Q^s)$ can be computed using Jacobi formula:

$$\partial_s \det(Q^s)|_{s=0} = a \operatorname{tr}(\partial_s Q^s\,(Q^s)^{-1})|_{s=0}.$$

An explicit expression of the derivatives of the map Q^s can be computed using Lemma 5.4. It follows that $b = \operatorname{tr}(\partial_s Q^s\,(Q^s)^{-1}) - \operatorname{tr}(K)$ and we obtain precisely the formula in the statement. $\square$

Before giving the explicit formula for $\mathrm{tr}(K)$ and the derivatives of the fundamental solution to Jacobi equation at $s = 0$ we need to make some notational remark and write down a formula for the second variation in the same spirit of eq. (2.7). We are working on the extended state space $M \times M$ with twice the number of variables of the original system and trivial dynamic on the first factor and separated boundary conditions. The left boundary condition manifold is the diagonal of $M \times M$ and the right one is our starting N. We apply the formula in eq. (2.7) to this particular system, we denote by $\tilde{Z}_t$ and $\tilde{Z}_i$ the matrices for the auxiliary problem, in general everything pertaining to it will be denoted with a tilde. Identifying $T_{(\lambda_0,\lambda_0)}T^*(M \times M)$ with $T_{\lambda_0}T^*M \times T_{\lambda_0}T^*M$ we have that:

$$\tilde{Z}_0 u_0 = \begin{pmatrix} Z_0 u_0 \\ Z_0 u_0 \end{pmatrix}, \quad \tilde{Z}_1 = \begin{pmatrix} Z_1^0 u_1 \\ Z_1^1 u_1 \end{pmatrix}, \quad \tilde{Z}_t = \begin{pmatrix} 0 \\ Z_t \end{pmatrix}.$$

We still work on the subspace $\mathcal{V} = \{(u_0, u_t, u_1) : \tilde{Z}_0 u_0 + \int_0^1 \tilde{Z}_t u_t + \tilde{Z}_1 u_1 \in \Pi\}$. However it is clear that the this equation implies that:

$$Z_0 u_0 + Z_1^0 u_1 \in \Pi_0 \quad \text{ and } \quad Z_0 u_0 + Z_1^1 u_1 + \int_0^1 Z_t u_t dt \in \Pi_0$$

It follows that control u_0 is completely determined by u_1. Moreover we can assume that $Z_0 u_0 = -Z_1^0 u_1$ since we are free to choose any system of coordinates and any trivialization of the tangent bundle of the manifolds Δ and N. Technically we are working with different scalar products on each of the copies of $T_{\lambda_0}T^*M$. However it is easy to see that on the space $\mathcal{V}$ only the sum of this metrics plays a role. We will denote it g_0. Now we are ready to state the following:

Lemma 5.3. *The second variation of the extended system, as a quadratic form, can be written as $\langle (I + K)u, u\rangle$ where K is the symmetric (on $\mathcal{V}$) compact operator given by:*

$$\begin{aligned} -\langle Ku, u\rangle = &\int_0^1 \int_0^t \sigma(Z_\tau u_\tau, Z_t u_t) d\tau dt - \sigma\left(Z_1^0 u_1, \int_0^1 Z_t u_t dt\right) \\ &+ \sigma\left(\int_0^1 Z_t u_t dt - Z_1^0 u_1, Z_1^1 u_1\right) + g_0(Z_1^0 u_1, Z_1^0 u_1) \\ &+ (\tilde{\Phi}^* g_1)(Z_1^1 u_1, Z_1^1 u_1). \end{aligned}$$

Moreover, define the following matrices:

$$\Gamma = \int_0^1 X_t X_t^* dt, \quad \Omega = \int_0^1 \int_0^t X_t Z_t^* J Z_\tau X_\tau d\tau dt.$$

Denote by pr_1 the projection onto $T_{(\lambda_0,\lambda_1)}A(N)$ and by π_*^i the differential of the natural projections $\pi^i : T^*M \to M$ relative to the $i-$th component, $i = 1, 2$. The trace of K has the following expression:

$$\begin{aligned} \mathrm{tr}(K) = &- \dim(N) + \mathrm{tr}[\pi_*^1 \tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_*(\tilde{Z}_0)] \\ &+ \mathrm{tr}\left[\Gamma^{-1}\left(\Omega + (\pi_*^2 - \pi_*^1)\tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_*\left(\int_0^1 \tilde{Z}_t Z_t^* J|_\Pi dt\right)\right)\right] \end{aligned}$$

Proof. The first part is a straightforward computation combining the expression obtained in Lemma 5.2 for the second variation with the observation concerning the structure of the maps Z_0 and Z_1 made before the statement and the choice of the Riemannian metrics.

Now, notice that the codimension of the space giving fixed endpoints variations $\mathcal{V}$ for the extended system in $\mathcal{H}$ is $2\dim(M)$. Moreover it is defined as the kernel of the linear functional:

$$\rho : (u_0, u_t, u_1) \mapsto \tilde{\pi}_*(\tilde{Z}_0 u_0 + \int_0^1 \tilde{Z}_t u_t dt + \tilde{Z}_1 u_1) \in T_{\pi(\lambda_0)}M \times T_{\pi(\lambda_0)}M.$$

It is straightforward to check that the following subspace is $2\dim(M)-$dimensional and transversal to $\ker\rho$:

$$\mathcal{V}' = \{(u_0, Z_t^* J\nu, 0) : \nu \in \Pi, u_0 \in \mathbb{R}^{\dim M}\}$$

The trace of K on the whole space splits as a sum of two pieces, the trace of $K|_{\mathcal{V}}$ and the trace of $K|_{\mathcal{V}'}$. We can then further simplify and compute separately the trace on $\mathcal{V}' \cap \{u_0 = 0\}$ and its complement $\mathcal{V}' \cap \{\nu = 0\}$. We are going to compute the trace of K on the whole space and then the trace of $K|_{\mathcal{V}'}$, determining in this way the value of $K|_{\mathcal{V}}$.

Consider $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Where

$$\mathcal{H}_1 = \{u : u = (0, u_t, 0)\}, \quad \mathcal{H}_2 = \{u : u = (u_0, 0, u_1)\}.$$

It is straightforward to check that $\mathcal{H}_1^\perp = \mathcal{H}_2$ for our class of metrics and that $\mathcal{H}_1 \equiv L^2([0,1], \mathbb{R}^k)$. Using eq. (2.7), the restrictions of the quadratic form $\hat{K}(u) = \langle u, Ku \rangle$ to each one of the former subspaces read:

$$\hat{K}|_{\mathcal{H}_1}(u) = \int_0^1 \int_0^t \sigma(Z_t u_t, Z_\tau u_\tau) dt d\tau, \quad \hat{K}|_{\mathcal{H}_2}(u) = \sigma(\tilde{Z}_0 u_0, \tilde{Z}_1 u_1) - \|u\|^2_{\mathcal{H}}.$$

The trace of the first quadratic form is zero (see [7][Theorem 2]) whereas the trace of the second part is just $-\dim(N)$. Thus we have that:

$$\mathrm{tr}(K|_{\mathcal{V}}) = -\dim(N) - \mathrm{tr}(K|_{\mathcal{V}'})$$

To compute the last piece we apply K to a control $u \in \mathcal{V}' \cap \{u_0 = 0\}$ using the explicit expression of the operator given in Lemma 5.2. Recall that pr_1 is projection onto the image of $\tilde{\Phi}_* \tilde{Z}_1$. It follows that:

$$\begin{aligned} K(u) &= \left(0, -Z_t^* J \int_0^t Z_\tau u_\tau d\tau, -L\, pr_1 \tilde{J}_1 \tilde{\Phi}_* \left(\int_0^1 \tilde{Z}_t u_t dt\right)\right) \\ &= \left(0, -Z_t^* J \int_0^t Z_\tau u_\tau d\tau, -\Lambda(u)\right). \end{aligned}$$

Now we write $K(u)$ in coordinates given by the splitting $\mathcal{V} \oplus \mathcal{V}'$. To do so we have first to consider $\rho \circ K(u)$. It is given by:

$$\begin{aligned} -\rho \circ K(u) &= \tilde{\pi}_* \left(\int_0^1 \tilde{Z}_t Z_t^* J \int_0^t Z_\tau Z_\tau^* J\nu + \tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_* (\int_0^1 \tilde{Z}_t Z_t^* J\nu dt)\right) \\ &= \begin{pmatrix} 0 \\ \Omega\nu \end{pmatrix} + \tilde{\pi}_* \tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_* (\int_0^1 \tilde{Z}_t Z_t^* J\nu dt). \end{aligned} \tag{5.20}$$

Set $\Gamma = \int_0^1 X_t^* X_t dt$, it easy to check that, if $u \in \mathcal{V}'$, then

$$\rho(u) = \begin{pmatrix} X_0 u_0 \\ \Gamma\nu + X_0 u_0 \end{pmatrix}$$

for an invertible matrix X_0 which, without loss of generality, can be taken to be identity. It follows that the projection on the first component of $\rho \circ K(u)$ is completely determined by the second term in eq. (5.20). Let us call π_*^i for $i = 1,2$ the projection on the $i-$th component. It follows that an element $(Z_0\hat{u}_0, Z_t^* J\hat{\nu}, 0) = \hat{u} \in \mathcal{V}'$ has the same projection as $K(u)$ if and only if:

$$\rho \circ (K(u) - \hat{u}) = 0 \iff \begin{cases} \hat{u}_0 & = X_0^{-1}\pi_*^1 \tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_*(\int_0^1 \tilde{Z}_t Z_t^* J\nu dt) \\ \hat{\nu} & = \Gamma^{-1}\left(\Omega\nu + (\pi_*^2 - \pi_*^1)\tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_*(\int_0^1 \tilde{Z}_t Z_t^* J\nu dt)\right) \end{cases}$$

In particular the restriction to $\mathcal{V}' \cap \{u_0 = 0\}$ is given by:

$$\nu \mapsto \Gamma^{-1}\left(\Omega + (\pi_*^2 - \pi_*^1)\tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_*(\int_0^1 \tilde{Z}_t Z_t^* J dt)\right)\nu.$$

A similar strategy applied to $\mathcal{V}' \cap \{\nu = 0\}$ tells us that the last contribution for the trace is given by the following map:

$$u_0 \mapsto X_0^{-1}\pi_*^1 \tilde{\Phi}_*^{-1} pr_1 \tilde{J}_1 \tilde{\Phi}_* \tilde{Z}_0 u_0.$$

It is worth pointing out that indeed the trace does not depend on X_0 and that the vector $\int_0^1 \tilde{Z}_t Z_t J\nu dt$ is the following:

$$\int_0^1 \tilde{Z}_t Z_t J\nu dt = \begin{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} \Theta\nu \\ \Gamma\nu \end{pmatrix} \end{pmatrix}$$

In particular if the boundary conditions are separated (i.e. $N = N_0 \times N_1$) the part of the trace coming from $\mathcal{V}' \cap \{u_0 = 0\}$ depends only on the projection onto $T_{\lambda_1}N_1$. □

Lemma 5.4. *The flow $\Phi_t^s|_{s=0}$ and its derivative $\partial_s \Phi_t^s|_{s=0}$ are given by:*

$$\Phi_t^s|_{s=0} = \begin{pmatrix} 1 & 0 \\ \int_0^t X_\tau X_\tau^* d\tau & 1 \end{pmatrix}, \quad \partial_s \Phi_t^s|_{s=0} = \begin{pmatrix} \int_0^t Y_\tau X_\tau^* d\tau & 0 \\ \int_0^t \int_0^\tau X_\tau Z_\tau^* J Z_r X_r^* dr d\tau & -\int_0^t X_\tau Y_\tau^* d\tau \end{pmatrix}$$

Proof. It is straightforward to check that $\Phi_1^s|_{s=0}$ and $\partial_s \Phi_1^s|_{s=0}$ solve the following Cauchy problems:

$$\begin{cases} \dot{\Phi}_t^0 = \begin{pmatrix} 0 & 0 \\ X_t X_t^* & 0 \end{pmatrix} \Phi_t^0, \\ \Phi_0^0 = Id. \end{cases} \qquad \begin{cases} \partial_s \dot{\Phi}_t^0 = \begin{pmatrix} 0 & 0 \\ X_t X_t^* & 0 \end{pmatrix} \partial_s \Phi_t^0 + \begin{pmatrix} Y_t X_t^* & 0 \\ 0 & -X_t Y_t^* \end{pmatrix} \Phi_t^0, \\ \partial_s \Phi_0^0 = 0. \end{cases}$$

Solving the ODE one obtains the formula in the statement. □

Appendix A

Symplectic Geometry and Maslov index

A.1 Linear symplectic geometry

Let Σ be a finite dimensional (real) vector space and σ a non degenerate skew-symmetric bilinear form on Σ. The couple (Σ, σ) is called symplectic vector space and σ the symplectic form. An example is given by $\mathbb{R}^{2n}$ with the standard symplectic form:

$$\sigma_{st}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1^T y_2 - x_2^T y_1, \quad x_i, y_i \in \mathbb{R}^n$$

Using the standard euclidean product on $\mathbb{R}^{2n}$, we can introduce a $2n \times 2n$ matrix J as follows and write the symplectic form as:

$$\sigma_{st}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rangle := \langle J \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rangle \tag{A.1}$$

It turns out that any symplectic subspace is isomorphic to the standard one just introduced, meaning that there exists a linear isomorphism $F : \Sigma \to \mathbb{R}^{2n}$ such that $\sigma_{st}(Fv, Fw) = \sigma(v, w)$ for all $v, w \in \Sigma$. Any such map is called (linear) *symplectomorphism*. The group of symplectomorphisms of $(\mathbb{R}^{2n}, \sigma_{st})$ is denoted by $Sp(2n, \mathbb{R})$. It is a non compact Lie group of dimension $n(2n + 1)$. It can be realized as a subgroup of $GL(2n, \mathbb{R})$, as the set:

$$Sp(2n, \mathbb{R}) = \{F \in GL(2n, \mathbb{R}) : FJF^T = J\}.$$

In a symplectic vector space we have a concept of *orthogonality* induced by the symplectic form. Given $V \subseteq \Sigma$ we denote by $V^\angle$ the following subspace:

$$V^\angle := \{v \in \Sigma : \sigma(v, w) = 0 \, \forall \, w \in V\}. \tag{A.2}$$

We can distinguish three classes of subspaces inside a symplectic space:

i) **Isotropic** if $V^\angle \supseteq V$;

ii) **Lagrangian** if $V^\angle = V$;

iii) **Coisotropic** if $V^\angle \subseteq V^\perp$.

We will be particularly interested in Lagrangian subspaces. They are characterized by the property of being the maximal subspaces on which σ restricts to the zero form and have always dimension n. It is well know that the group $Sp(2n, \mathbb{R})$ acts transitively on couples of transversal Lagrangian subspace (see for instance [27, Theorem 1.15]). This is not the case for triples. This fact will be central in the definition of Maslov index. In the sequel we will also need the following fact. Fix L_0 and L_1 Lagrangian subspaces then one can always find coordinates in which the symplectic structure becomes the standard one, L_0 is mapped to the span of the first n coordinates and L_1 to the span of the last n.

The collection of all Lagrangian subspaces is called Lagrange Grassmannian and denoted by:

$$Lag(\Sigma) = \{V \subset \Sigma : \sigma(v, w) = 0 \,\forall\, v, w \in V\}.$$

It is a smooth, closed $n(n+1)$ algebraic manifold. Its geometry will be discussed in the subsequent sections.

Even if all symplectic spaces are equivalent to standard one, we list here some common constructions to get new symplectic spaces.

Example A.1. Take a symplectic space (Σ, σ) and define $\Sigma' = \Sigma \oplus \Sigma$ with symplectic form given by $(-\sigma) \oplus \sigma$:

$$((-\sigma) \oplus \sigma)\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = \sigma(v_2, w_2) - \sigma(v_1, w_1).$$

Σ' is a symplectic space and more importantly graphs of symplectic maps are Lagrangian subspaces.

Example A.2. Take a vector space V and its dual V^*. Define the vector space $\Sigma = V \oplus V^*$ with symplectic form:

$$\sigma\left(\begin{pmatrix} \lambda_1 \\ v_1 \end{pmatrix}, \begin{pmatrix} \lambda_2 \\ v_2 \end{pmatrix}\right) = \lambda_1(v_2) - \lambda_2(v_1).$$

One can check that σ is non degenerate. We have a distinguished class of Lagrangian subspaces sitting inside $Lag(\Sigma)$. Take $W \subseteq V$ and define its annihilator to be:

$$Ann(W) = \{\lambda \in V^* : \lambda(w) = 0 \,\forall\, w \in W\}.$$

It is straightforward that the subspaces of the form $W \oplus Ann(W)$ have dimension n and are Lagrangian.

A.2 Lagrange Grassmannians and intersection indices

In this section we collect some information about the geometry of Lagrange Grassmannian and curves in it (i.e. $1-$parameter families of Lagrange subspaces).

First of all let us introduce an atlas on $Lag(\Sigma)$ made up of affine charts. Take a couple of transversal Lagrangian subspaces L_0 and L_1. By what stated in the previous section we can always assume $\Sigma \approx \mathbb{R}^{2n}$ with the identifications:

$$L_0 \approx \{(y, 0) : y \in \mathbb{R}^n\}, \quad L_1 \approx \{(0, x) : x \in \mathbb{R}^n\}.$$

For any symmetric $n \times n$ matrix S, we can consider the subspace $L_S := \{(Sx, x) : x \in \mathbb{R}^n\}$. It is straightforward to check that L_S is Lagrangian and transversal to L_0. This correspondence gives a map between the space of symmetric matrices and an open set of the Lagrange Grassmannian. We call this set $L_0^{\pitchfork}$ since it coincides with the collection of all subspaces transversal to L_0.

The complementary set to $L_0^{\pitchfork}$, which we call the train of L_0 and we denote by $\mathcal{M}_{L_0}$, is a singular hypersurface in $Lag(\Sigma)$. One can check using the charts defined above that it coincides locally with the set $\{S \in Mat_n(\mathbb{R}), S = S^T, \text{rank}(S) < n\}$. In particular the rank induces a stratification of $\mathcal{M}_{L_0}$. The smooth stratum corresponds to $\{\text{rank}(S) = n-1\}$ whereas the singular ones have codimension at least 3.

It is possible to endow $\mathcal{M}_{L_0}$ with a co-orientation, in plain words this means that at each point we can choose consistently an outer normal to $\mathcal{M}_{L_0}$. This is because of the following interpretation of tangent vectors to $Lag(\Sigma)$. Take a curve of Lagrangians L_t and, for each t define the following quadratic form on L_t:

$$Q(l_t) = \sigma(l_t, \dot{l}_t), \quad l_t \in L_t$$

Here l_t denotes a curve inside L_t and $\dot{l}_t$. The formula above defines a quadratic form which corresponds to the velocity of L_t. To see this correspondence we fix one of the chart mentioned above and write $L_t = \{(S_t x, x) : x \in \mathbb{R}^{2n}\}$. It is straight forward to compute Q on a curve of the type $t \mapsto (S_t x, x)$ and see that $Q((S_t x, x)) = \langle x, \dot{S}_t x\rangle$.

This in particular allows to define a co-orientation on (the smooth part of) $\mathcal{M}_{L_0}$ as follows. Suppose that $L_t \cap L_0 = \text{span}\{l\}$, we say that L_t crosses positively if $\dot{L}_t(l) > 0$. The facts that $\mathcal{M}_{L_0}$ has singularity in codimension grater than 2 and it is co-orientable, allows us to speak about general position and intersection number of smooth curves with $\mathcal{M}_{L_0}$.

Definition A.1 (Maslov Index of a smooth curve)**.** Fix a Lagrange subspace L_0 and a smooth curve L_t of Lagrangian subspaces. Assume that the endpoints of L_t are transversal to L_0. Maslov Index is defined as the intersection number of L_t with $\mathcal{M}_{L_0}$. If L_t is in general position it holds:

$$\text{Mi}_{L_0}(L_t) = \sum_{t: L_t \cap L_0 = (l)} \text{sgn}(\dot{L}(l)).$$

Luckily enough it is not necessary to put curves in general positions to compute their Maslov index. Using the affine charts introduced above one can prove the following (see [11][Chapter 1]):

Proposition A.1. *Assume that the curve L_t, $t \in [t_0, t_1]$ lives inside a chart of the type $L_2^{\pitchfork}$ and that $L_2 \cap L_0 = (0)$. Then, in coordinates given by L_0 and L_2, there exists a family of symmetric matrices such that $L_t = \{(S_t x, x) : x \in \mathbb{R}^n\}$. The Maslov index of L_t is given by:*

$$\text{Mi}_{L_0}(L_t) = \text{ind}^- S_{t_0} - \text{ind}^- S_{t_1}.$$

Here $\text{ind}^- S_t$ *denotes the number of negative eigenvalues (or Morse index) of* S_t.

In particular given a curve L_t with $t \in [t_0, t_1]$ we can find a partition $\{s_0 < \dots < s_m\}$ of $[t_0, t_1]$ and finite number of Lagrangian subspaces transversal to L_0 and $L_t|_{[s_i, s_{i+1}]}$ and present the Maslov index of a curve as the sum of local contribution as in the proposition. To the reader familiar with a little algebraic topology this construction may look very similar to the one for an *angle function* in $\mathbb{C}\setminus\{0\}$. Indeed both the spaces have fundamental group isomorphic to $\mathbb{Z}$ (see the remark below) and, when considering closed loops, the Maslov index gives a morphism between $\pi_1(Lag(\Sigma))$ and the integers.

We are going now to list some of the properties of Maslov index we will use in the sequel. For proofs and detailed discussion one can look in [27, 14, 6, 11]. First we will need the following definition:

Definition A.2. A curve L_t is (strictly) positive if its velocity satisfies $\dot{L}_t \geq 0$ (respectively > 0). We say that it is (strictly) negative if $\dot{L}_t \leq 0$ (respectively < 0).

Proposition A.2. *i) We say that two curves L_t and L'_t defined on $[t_0, t_1]$ if there exists a smooth family $\Lambda_{t,s}$ such that $\Lambda_{t,0} = L_t$ and $\Lambda_{t,1} = L'_t$ and $\Lambda_{1,s}$ and $\Lambda_{0,s}$ do not intersect $\mathcal{M}_{L_0}$ for $s, t \in [t_0, t_1]$. If L_t and L'_t are homotopic then:*

$$\mathrm{Mi}_{L_0}(L_t) = \mathrm{Mi}_{L_0}(L'_t).$$

ii) If L_0 and L'_0 are two Lagrangian subspaces transverse to both L_{t_0} and L_{t_1}, the Maslov index satisfies:

$$|\mathrm{Mi}_{L_0}(L_t) - \mathrm{Mi}_{L'_0}(L_t)| \leq n.$$

iii) If L_t is a closed curve then for any L_0 and L'_0 we have:

$$\mathrm{Mi}_{L_0}(L_t) = \mathrm{Mi}_{L'_0}(L_t).$$

iv) If L_t is a strictly positive curve with endpoints transversal to L_0 then:

$$\mathrm{Mi}_{L_0}(L_t) = \sum_{t \in [t_0, t_1]} \dim(L_t \cap L_0).$$

Remark A.1. It turns out that Lagrange Grassmannians are homogeneous space. In fact one can show that the group $U(n)$, which we identify as a subgroup of $GL(2n, \mathbb{R})$, acts transitively on $Lag(\mathbb{R}^{2n})$. The stabilizer of this action is given by a subgroup isomorphic to $O(n)$ and thus $Lag(\mathbb{R}^{2n}) \approx U(n)/O(n)$. This implies in particular that $Lag(\Sigma)$ is not simply connected and its fundamental group is isomorphic to $\mathbb{Z}$.

Maslov index and crossing form

In this section we introduce a quadratic form which we will need in the following, the *Maslov form*. We will also relate it with the Maslov index of path introduce in the previous section and then prove some properties of the Maslov form needed in the sequel.

Definition A.3 (Maslov index)**.** Take three Lagrangian subspaces L_0, L_1, L_2. Consider the isotropic subspace $L'_1 := L_1 \cap (L_0 + L_2)$, if $l_1 \in L'_1$ then $l_1 = l_0 + l_2$ with $l_i \in L_i$. The following quadratic form is called the Maslov form of the triple (L_0, L_1, L_2):

$$m(l'_1) = \sigma(l_0, l_2).$$

If we want to be explicit about which Lagrangian subspaces are being used ,instead of just m, we will write $m(L_0, L_1, L_2)$.

The numbers $\mathrm{ind}^+ m$, $\mathrm{ind}^- m$ and $sg\, m$ and $\dim\ker m$ are invariants of the triple (L_0, L_1, L_2). The *Kashiwara index* is the signature of the Maslov form:

$$\tau(L_0, L_1, L_2) = sg\, m = \mathrm{ind}^+ m - \mathrm{ind}^- m.$$

The *negative Maslov index* is defined as

$$i(L_0, L_1, L_2) = \operatorname{ind}^- m. \tag{A.3}$$

Example A.3. Suppose L_0 and L_2 are transversal. We can identify the symplectic space with the standard one $(\mathbb{R}^{2n}, \sigma)$ as given in eq. (A.1). Since any couple of Lagrangian subspaces can be mapped into each other, we can find a symplectomorphism which simultaneously maps L_0 to B and L_2 to Π.

Any L_1 can be represented as $L_1 = \{Aq + Cp = 0, q \in B, p \in \Pi\}$ where $AC^* = CA^*$ and $rank[A, C] = n$. If A or C is invertible then the matrix expression of the Maslov form is given by $-A^{-1}C$ or $C^{-1}A$ respectively, which have the same signature of $\mp AC^*$.

The Kashiwara index and the Maslov index have the following properties:

i) (*Alternating*) $\tau(L_{s(0)}, L_{s(1)}, L_{s(2)}) = (-1)^{sg(s)}\tau(L_0, L_1, L_2)$, where s is a permutation.

ii) (*Cocycle property*, [27, Theorem 1.32])

$$\tau(L_0, L_1, L_2) - \tau(L_1, L_2, L_3) + \tau(L_0, L_2, L_3) - \tau(L_0, L_1, L_3) = 0. \tag{A.4}$$

iii) (*Relation between the negative index*, [6, Lemma 5])

$$\begin{aligned}\tau(L_0, L_1, L_2) &= -2\, i(L_0, L_1, L_2) + \dim(L_1 \cap (L_0 + L_2)) - \dim\ker m\\ &= -2\, i(L_0, L_1, L_2) + n - \dim(L_0 \cap L_2) - \dim(L_0 \cap L_1)\\ &\quad - \dim(L_1 \cap L_2) + 2\dim(L_0 \cap L_1 \cap L_2).\end{aligned} \tag{A.5}$$

iv) (*Symplectic reduction*) If $V \subseteq L_0 \cap L_2$ is an isotropic subspace we can consider $L^V := (L \cap V^\perp + V)/V$ which is a Lagrangian subspace of the reduced space. It holds that:

$$i(L_0, L_1, L_2) = i(L_0^V, L_1^V, L_2^V) \tag{A.6}$$

Suppose now that L_t is a smooth curve inside $Lag(\Sigma)$ and L_0 a fixed Lagrangian subspace. Suppose that $V_{t'} = L_{t'} \cap L_0 \neq (0)$ for some $t' \in (t_0, t_1)$ and that $\dot{L}_{t'}|_{V_{t'}}$ is a non degenerate quadratic form. The quadratic form $\dot{L}_{t'}|_{V_{t'}}$ is sometimes called *crossing form*. We have the following proposition (see for instance [30, 11, 27] for a proof and compare with proposition A.1):

Proposition A.3. *Assume that L_0' is transversal to L_0 and $L_t|_{[t'-\epsilon, t+\epsilon]}$. Under the hypothesis above, for $\epsilon > 0$ small enough:*

$$\begin{aligned}2\operatorname{Mi}_{L_0}(L_t|_{[t'-\epsilon, t'+\epsilon]}) &= \tau(L_0', L_{t'-\epsilon}, L_0) - \tau(L_0', L_{t'+\epsilon}, L_0)\\ &= sg\, \dot{L}_{t'}|_{L_{t'} \cap L_0}.\end{aligned}$$

In particular one recovers point *iv*) of proposition A.2 assuming just that the restriction of the velocity of L_t is positive *on the intersection* $L_t \cap L_0$.

Unlike the Kashiwara index, the Maslov index satisfies only a *generic* cocycle property. This is shown in the next lemma which will be needed later on. The defect of being a cocycle is measured by the intersections of the four Lagrangian subspaces one considers.

Lemma A.1. *The following formulas hold:*

$$\sum_{i=0}^{3}(-1)^{i+1}i\big(L_0,\dots,\hat{L}_i,\dots L_3\big) = \dim(L_1\cap L_3) - \dim(L_0\cap L_2) + \\ + \sum_{i=0}^{3}(-1)^{i+1}\dim(L_0\cap\cdots\cap\hat{L}_i\cap\cdots\cap L_3). \tag{A.7}$$

$$i\big(L_0,L_1,L_2\big) = i\big(L_1,L_2,L_0\big) = i\big(L_2,L_0,L_1\big). \tag{A.8}$$

Proof. The proof is just a computation using the cocycle identity for the signature in eq. (A.4).

We use formula (A.5) to get a relation between the signature and the index for each term on the left of (A.7). The part involving intersections of pairs of Lagrangian spaces, after the summation, gives the contribution

$$2\Big(\dim(L_1\cap L_3) - \dim(L_0\cap L_2)\Big).$$

The triple intersection do not simplify and thus their coboundary appears. Sometimes it is useful to think about this remainder as the difference of the dimensions of two quotient spaces. $L_1\cap L_3$ in which we factor out the space $L_1\cap L_3\cap L_0 + L_1\cap L_3\cap L_0$ and $L_0\cap L_2$ quotient by $L_0\cap L_2\cap L_1 + L_0\cap L_2\cap L_3$.

Next apply the first part with $L_3 = L_0$. The right hand side of the formula is then zero. Moreover the terms $i\big(L_0,L_1,L_0\big)$ and $i\big(L_0,L_0,L_2\big)$ are zero since the Maslov form is zero.

It follows that $i\big(L_0,L_1,L_2\big) = i\big(L_1,L_2,L_0\big)$, i.e. we obtain the invariance under cyclic permutations. □

A.2.1 The complex picture

In this section we will sketch the main features of the complex version of Maslov index and symplectic geometry just introduced. We work almost always with real coefficients but at some points we need to switch to complex ones, mainly for spectral issues.

Take as vector space $\mathbb{C}^{2n} \approx \mathbb{R}^{2n}\otimes\mathbb{C}$. Consider the complex version of the symplectic form:

$$\sigma_{\mathbb{C}}(X,Y) = \sigma(\bar{X},Y), \quad X,Y\in\mathbb{C}^{2n}.$$

Clearly $\sigma_{\mathbb{C}}$ remains a non degenerate bilinear form. We can speak, as in the real case of, Lagrangian, isotropic and co-isotropic subspaces. Here we list some of the properties of the $\sigma_{\mathbb{C}}$ and complex Lagrange Grassmannian:

i) (*Darboux basis*) Since $\sigma_{\mathbb{C}}$ is non degenerate, every time two Lagrangian subspaces L_0, L_1 are considered, there exists a basis in which $\sigma_{\mathbb{C}}$ has the standard form.

ii) (*Grassmannian of Lagrangian subspaces*) In the real case the Lagrange Grassmannian is a homogeneous space diffeomorphic to $U(n)/O(n)$. It turns out that the *complex* one is diffeomorphic to $U(n)$ (and thus still a real manifold). We can diagonalize the symplectic form obtaining:

$$\frac{1}{2}\begin{pmatrix} i & -1\\ 1 & -i\end{pmatrix}\begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix} -i & 1\\ -1 & i\end{pmatrix} = \begin{pmatrix} i & 0\\ 0 & -i\end{pmatrix}.$$

Thus we have two subspaces on which $\sigma_{\mathbb{C}}$ is non degenerate, the eigenspace V_i relative to i and V_{-i}, the one relative to $-i$. It is thus clear that if V is Lagrangian, V must be transversal to both the eigenspaces. So it can always be represented as a graph of an invertible linear operator from $V_i \to V_{-i}$ (or vice versa). It remains to check what kind of linear maps are allowed. Using again the coordinates in which $\sigma_{\mathbb{C}}$ is diagonal we get:

$$\sigma_{\mathbb{C}}\Big(\begin{pmatrix} x \\ Rx \end{pmatrix}, \begin{pmatrix} y \\ Ry \end{pmatrix} \Big) = i\langle \bar{x}, y\rangle - i\langle R^*\bar{R}\bar{x}, y\rangle = i\langle (1 - R^*\bar{R})\bar{x}, y\rangle.$$

Since we need this quantity to be zero for any $x, y \in \mathbb{C}^n$ we get $\bar{R}R^* = 1$ and thus $R \in U(n)$. It follows that the complex Grassmannian is diffeomorphic to $U(n)$.

iii) (*Atlas for the Lagrange Grassmannian*) Take two transversal subspaces L_0, L_1. Using Darboux coordinates, we can build an affine chart as in the real case (compare with appendix A.2). This time, though, we consider Hermitian matrices. The subspaces $V_S = \{(x, Sx) : S = \bar{S}^*, x \in L_0\}$ are Lagrangian subspaces.

iv) (*Inclusion of the real version*)Notice that if V is a real Lagrangian subspace, then $V \otimes \mathbb{C}$ is Lagrangian with respect to the new symplectic form. In particular the real Lagrange Grassmannian embeds in the complex one.

Take now three Lagrangian subspaces, the Maslov form is still well defined. Suppose that $\lambda_1 \in L_1 \cap (L_0 + L_2)$:

$$m(\lambda_1) = \sigma_{\mathbb{C}}(\lambda_0, \lambda_2) = \sigma(\bar{\lambda}_0, \lambda_2).$$

Lemma A.2. *The Maslov form is a Hermitian form.*

Proof. Notice that $\overline{m(\lambda_1)} = \sigma(\lambda_0, \bar{\lambda}_2) = m(\bar{\lambda}_1)$. We have to show that $m(\lambda_1) = m(\bar{\lambda}_1)$ but this follows from the fact the subspaces are Lagrangian.

$$\begin{aligned} m(\lambda_1) - m(\bar{\lambda}_1) &= \sigma(\bar{\lambda}_0, \lambda_2) - \sigma(\lambda_0, \bar{\lambda}_2) = \sigma(\bar{\lambda}_0 + \bar{\lambda}_2, \lambda_2) - \sigma(\lambda_0, \bar{\lambda}_2 + \bar{\lambda}_0) \\ &= \sigma(\bar{\lambda}_1, \lambda_2) - \sigma(\lambda_0, \bar{\lambda}_1) = \sigma(\bar{\lambda}_1, \lambda_1) = 0. \end{aligned}$$

This means that the quadratic form is real and thus m is Hermitian. □

Thus the eigenvalues of m are real and the index and the signature are well defined exactly as in the real case. More importantly, if we take Lagrangian subspaces of the form $L_0 \otimes \mathbb{C}$,$L_1 \otimes \mathbb{C}$ and $L_2 \otimes \mathbb{C}$ the real and complex invariants coincide! All the properties of Maslov and Kashiwara index listed in the previous section extend in a natural way to this context, with real dimensions replaced by complex ones. The proof of [27, Theorem 1.32] works in the Hermitian case as well, since our quadratic forms are all real by Lemma A.2. It suffice to substitute the word *symmetric* with the word *Hermitian* in the proofs.

A.3 Symplectic Manifolds

Let M be a smooth manifold, if there exists a non degenerate closed 2−form σ then (M, σ) is called a symplectic manifold. In particular this means that each tangent space (T_pM, σ_p) is a symplectic vector space for every $p \in M$.

An example of symplectic manifold we will encounter very often is the cotangent bundle of a smooth manifold, denoted by T^*M. The symplectic form is given by the differential of the so called *tautological* (or *Liouville*) $1-$form s. It is defined as follows:

$$s_\lambda(X) = \lambda(\pi_* X), \quad \forall\, \lambda \in T^*M, X \in T_\lambda T^*X.$$

Here $\pi : T^*M \to M$ denotes the natural projection and π_* its differential. In local coordinates (p, q) ($q = (q_i)_i$ being local coordinates on M and $p = \sum_i p_i dq_i$ the ones on the fibre) one can check that $ds = dp \wedge dq$ and thus defines a non degenerate $2-$form. Notice that with this choices we are essentially identifying a chart of T^*M with an open set of the model space $(\mathbb{R}^{2n}, \sigma_{st})$. This turns out to hold for any symplectic manifold as the next theorem shows.

Theorem A.1 (Darboux). *Let (M, σ) be a symplectic manifold and $m \in M$. Then there exists an open neighbourhood U of m and coordinates (p, q) such that $\sigma = dp \wedge dq$ on U.*

Suppose now that $F : (M, \sigma) \to (M', \sigma')$ is a diffeomorphism, if $F^*\sigma' = \sigma$ we say that F is a *symplectomorphism.* A natural way to obtain diffeomorphisms is trough flows. Given a (complete) vector field X one obtains a family of diffeomorphisms Φ_t by solving the ODE system

$$\begin{aligned} \dot{\Phi}_t &= X(\Phi_t), \\ \Phi_0 &= Id. \end{aligned}$$

In a similar way one can produce symplectomorphisms using special classes of vector fields: *Hamiltonian* and *symplectic* fields. A vector field X, is *Hamiltonian* if there is a smooth function H such that $dH(Y) = \sigma(Y, X)$ for all smooth vector fields Y. H is called *Hamiltonian* function and X is often denoted by $\vec{H}$. A vector field for which we can find a Hamiltonian only locally, i.e. in a neighbourhood of every point, is called *symplectic.* The flow of Hamiltonian and symplectic vector fields is always a one parameter group of symplectomorphisms.

As for the linear case we can give the notions of Lagrangian, isotropic and co-isotropic submanifold. We say that N is a Lagrangian (resp. isotropic or co-isotropic) submanifold if T_nN is a Lagrangian (resp. isotropic or co-isotropic) subspace of T_nM for any $n \in N$. Curves are obvious example of isotropic manifold, if the ambient space is a cotangent bundle then we have two important classes of examples of Lagrangian submanifolds.

Example A.4. Assume that $f : M \mapsto R$ is a smooth function, then $df : M \to T^*M$ is a section of T^*M and gives and embedding of M into T^*M. It is straightforward to check, in local coordinates, that the tangent space to $df(M)$ can be identify with the graph of the Hessian of f, which is obviously Lagrangian.

Example A.5. Let $\lambda \in T^*M$ and consider the following subspace, which we usually call *fibre*:

$$\Pi_\lambda := \{\xi \in T_\lambda(T^*M) : \pi_*(\xi) = 0\} = \ker \pi_*. \tag{A.9}$$

Using the linear structure of T^*M it is easy to see that this is the tangent space to $\ker \pi$ at λ.

Example A.6. Assume that $N \subseteq M$ is a submanifold and let $n \in N$. We define the following subspace of T_n^*M:

$$Ann(T_nN) := \{\lambda \in T_n^*M : \lambda(X) = 0, \forall X \in T_nN\}.$$

By example A.2, $Ann(T_nN) \oplus T_nN$ is Lagrangian vector space, one can check in local coordinates that this is the tangent space to:

$$Ann(N) := \cup_{n \in N} Ann(T_nN).$$

For further reference we give also the following definition:

$$A_\lambda(N) := T_\lambda Ann(N). \tag{A.10}$$

where we omit the dependence on the point λ when there is no confusion. Be careful, however, that we often consider manifold $M \times M$ and their cotangent bundle $T^*M \times T^*M$ with symplectic form $(-\sigma) \oplus \sigma$, as in example A.1. This is not the tautological form, however the two structure are related by the following symplectomorphism:

$$S : (\lambda, \eta) \mapsto (-\lambda, \eta) \in T^*M \times T^*M.$$

It is easy to check that the $S^*((-\sigma) \oplus \sigma)$ is the tautological form $\sigma \oplus \sigma$. Often we will write $A(N)$ instead of $S\,A(N)$ and just specify which symplectic form we are using.

We now prove the following Lemma about the Maslov form we will use in the sequel.

Lemma A.3. *Given the standard symplectic structure $(\Sigma, \sigma) = (T_\lambda(T^*M), ds)$ and three submanifolds $N_0, N_1, N_2 \subset M$. Assume that:*

$$\lambda \in Ann(N_0) \cap Ann(N_1) \cap Ann(N_2)$$

and $N_1 \subseteq N_0$ (or $N_1 \subseteq N_2$) , then following formula holds

$$m\big(A_\lambda(N_0), A_\lambda(N_1), A_\lambda(N_2)\big) \equiv 0.$$

The same holds if $M = M' \times M'$ and T^*M is endowed with the form $(-\sigma') \oplus \sigma'$. In this case $A(N_i)$ are modified accordingly as explained after eq. (A.10).

Proof. Let $L_0 = T_\lambda A(N_0)$, $L_1 = T_\lambda A(N_1)$ and $L_2 = T_\lambda A(N_2)$. Fix some coordinates in a neighbourhood of λ such that ds_λ is the standard form on $\mathbb{R}^{2n} \simeq T_\lambda(T^*M)$. The subspace $L_0 + L_2$ is the space:

$$\begin{pmatrix} \nu_0 \\ X_0 \end{pmatrix} + \begin{pmatrix} \nu_2 \\ X_2 \end{pmatrix}, \qquad X_i \in T_{\pi(\lambda)}N_i, \qquad \nu_i(T_{\pi(\lambda)}N_i) = 0,$$

for $i = 0, 2$. Since the sum above should lie in $L_1 \cap (L_0 + L_2)$, we have that $X_0 + X_2 = X_1 \in T_{\pi(\lambda)}N_1$ and $\nu_0 + \nu_2 = \nu_1$, with ν_1 such that $\nu_1(T_{\pi(\lambda)}N_1) = 0$. If we compute now the Maslov form, we get:

$$\left\langle J \begin{pmatrix} \nu_0 \\ X_0 \end{pmatrix}, \begin{pmatrix} \nu_2 \\ X_2 \end{pmatrix} \right\rangle = \langle \nu_0, X_2 \rangle - \langle \nu_2, X_0 \rangle.$$

Suppose without loss of generality that $N_1 \subseteq N_0$. The equation $X_0 + X_2 = X_1$ implies that $X_2 = X_1 - X_0 \in T_{\pi(\lambda)}N_0$ and thus $\langle \nu_0, X_2 \rangle = 0$. Therefore the quadratic form is the zero form since:

$$\langle \nu_2, X_0 \rangle = \langle \nu_2, X_0 + X_2 \rangle = \langle \nu_2 + \nu_0, X_0 + X_2 \rangle = \langle \nu_1, X_1 \rangle = 0.$$

For the second part, we work on the cotangent bundle of $M = M' \times M'$ which is isomorphic to $T^*M' \times T^*M'$. Label the coordinates as (λ_0, λ_1), call the standard form on T^*M', σ' and consider the following diffeomorphism:

$$S : (\lambda_0, \lambda_1) \mapsto (-\lambda_0, \lambda_1).$$

It is straightforward to check that $S^*(\sigma' \oplus \sigma') = (-\sigma') \oplus \sigma'$ and S maps $A(N_i)$ as given in eq. (A.10) to $SA(N_i)$. Since Maslov index is invariant with respect to the action of symplectomorphisms, the statement follows.

$\square$

Appendix B

Control systems

In this section we collect some basic facts about control system and the endpoint map, mainly to set notation and as reference for the other chapters.

Suppose that M is a smooth manifold, U is a subset of $\mathbb{R}^k$ and f_u^t a family of smooth, complete possibly non autonomous vector fields. We assume that they are measurable and locally bounded in t if they explicitly depend on time. We say that (M, f_u^t) is *control system*, we call M *state space* and U the *control parameter space*. Any point $q_0 \in M$ is called state and any L^∞ function $u : [0,1] \to U$ *control*. To any couple $(q_0, u(t))$ we can associate a Cauchy problem:

$$\begin{cases} \dot{q} = f_{u(t)}(q), \\ q(0) = q_0. \end{cases} \tag{B.1}$$

We usually work with $[0,1]$ for notational convenience, but any interval $[0,\tau]$ would do. For fixed $q_0 \in M$, we say that a control is *admissible* if the corresponding Cauchy problem has solution defined on $[0,1]$. Any solution to the Cauchy problem above with u admissible is called *admissible* trajectory. We will denoted any such curve by $q_u(t)$ to underline the dependence on the control. The following proposition si proved in [14]:

Proposition B.1. *The space of admissible trajectories is a smooth Banach manifold modelled on the space $L^\infty([0,1], \mathbb{R}^k) \times \mathbb{R}^n$.*

We define now one of the main tools of geometric control theory, the Endpoint map. Take a couple (q_0, u) of state and admissible control, for $t \in [0,1]$ we define the *endpoint map* at time t:

$$E_{q_0}^t(u) = q_u(t), \quad q_u(\cdot) \text{ solution of eq. (B.1)}.$$

It turns out that $E_{q_0}^t$ is a smooth map from the space of admissible controls to the state space M. In particular, if we fix the initial condition and look at the level sets $(E_{q_0}^t)^{-1}(q)$ for $q \in M$, and assume that the differential of the endpoint map is surjective, they are smooth submanifolds of the space of admissible controls.

Suppose now that you are given a smooth, in the space and control variable, function $\varphi(t, u, q)$, measurable in t and locally bounded. We can define an integral functional on the space of admissible controls as follows:

$$\min_{u \in (E_{q_0}^t)^{-1}(q)} \mathcal{J}(u) = \min_{u \in (E_{q_0}^t)^{-1}(q)} \int_0^1 \varphi(t, u(t), q_u(t))dt. \tag{B.2}$$

Minimizing the functional in eq. (B.2) and finding the control u which realizes the minimum is called an *optimal control* problem.

Remark B.1. The functional in eq. (B.2) correspond to variational problems with Dirichlet boundary conditions, i.e. we are minimizing the cost among curves with the same end-points. We can generalize it slightly considering couples of initial condition and control (q_0, u) and requiring that $(q_0, E^t_{q_0}(u))$ is contained in a fixed submanifold $N \subseteq M \times M$, the *boundary conditions*.

The first step to approach it, is to find the critical points of the functional $\mathcal{J}$ on the space of admissible controls. They are characterized by the so called first order necessary optimality condition and can be summarize in the Theorem we are going to state now, known as Pontryagin Maximum Principle. Define the following family of functions on $T^*M \times \mathbb{R}$:

$$h^t_u(\lambda, \nu) = \langle \lambda, f_u \rangle + \nu\varphi(t, u, q).$$

Theorem B.1 (PMP). *If $\tilde{u}_t$ is a critical point of the functional $\mathcal{J}$ and $\tilde{q}_t$ is the relative admissible trajectory, then there exists a Lipschitz curve $\lambda_t : [0,1] \to T^*M$ and a number $\nu \in \mathbb{R}$ such that:*

i) $\tilde{q}_t = \pi\lambda_t$ *for all* $t \in [0,1]$*;*

ii) $(\lambda_t, \nu) \neq 0$ *for a.e.* $t \in [0,1]$*;*

iii) $\nu \leq 0$*;*

iv) $\dot{\lambda}_t = \vec{h}^t_{\tilde{u}(t)}(\lambda_t, \nu)$ *for a.e.* $t \in [0,1]$*;*

v) $h^t_{\tilde{u}(t)}(\lambda_t, \nu) = \max_{u \in U} h^t_u(\lambda_t, \nu)$ *for a.e.* $t \in [0,1]$.

Moreover if we are working with boundary conditions $N \subseteq M \times M$ (see remark B.1) we have the additional condition:

vi) $(-\lambda_0, \lambda_1) \in Ann(T_{(\tilde{q}_0, \tilde{q}_1)}N)$.

There are essentially two possibility for the value of the parameter ν, it can be either 0 or -1 after re-parametrizing λ_t.

Definition B.1. Any curve satisfying the conditions of Pontryagin Maximum Principle with $\nu = -1$ is called *normal* extremal. If $\nu = 0$ we will call it *abnormal* extremal.

B.1 Optimal control problems via Lagrange multipliers

The purpose of this section is to rephrase the problem in eq. (B.2) as a constrained minimization problem with the aid of *Lagrange multipliers* rule. If $f : \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}^m$ are smooth functions, the classical result from calculus says that, looking for critical points of f restricted to the zero level set of g, if the latter is *regular*, is equivalent to look for *free* critical points on $\mathbb{R}^d \times \mathbb{R}^m$ of the function $f(x, \lambda) = f(x) - \langle \lambda, g(x) \rangle$. In other words one has to solve the equations $g(x) = 0$, and $\nabla f(x) - \lambda \mathrm{Jac}(g)(x) = 0$ in λ and x if $\mathrm{Jac}(g)(x)$ has full rank along $\{g(x) = 0\}$.

We need a slight generalization of this principle, which allows the function g to have singularities. It says that finding critical points of f constrained to $\{g(x) = 0\}$ is equivalent to solving the system $\nu\nabla f(x) - \lambda \mathrm{Jac}(g)(x) = 0$ and $g(x) = 0$ in x, λ and ν. This is is the same as the following condition: define that map $F = (f, g)$, finding critical points of f on the set $\{g(x) = 0\}$ is equivalent to finding critical points of F.

It turns out that as long as the codomain of g is finite dimensional, the same principle holds. First order conditions for the optimal control problem in eq. (B.2) are equivalent to finding critical points of $\mathcal{J}$ restricted to the level sets of the endpoint map and thus equivalent to critical points of the following extended Endpoint map:

$$F^t_{q_0}(u) = \begin{pmatrix} \int_0^t \varphi(\tau, u(\tau), q_u(\tau))d\tau \\ E^t_{q_0}(u) \end{pmatrix} = \begin{pmatrix} \mathcal{J}^t_{q_0}(u) \\ E^t_{q_0}(u) \end{pmatrix} \tag{B.3}$$

In particular optimal control problems in eq. (B.2) with Dirichlet boundary conditions (i.e. minimizing among Lipschitz trajectory joining q_0 and q_1) are equivalent to the following system:

$$\begin{cases} E^t_{q_0}(u) = q_1, \\ \lambda_t \, dE^t_{q_0}(u) - \nu \, d\mathcal{J}^t_{q_0}(u) = 0 \end{cases}$$

We are going now derive integral expression for the first and second differential of the endpoint map $E^t_{q_0}$ (and of $F^t_{q_0}$). However to do so we need to introduce some tools of Chronological Calculus. Details can be found in [12].

B.1.1 Chronological Calculus

The scope of Chronological calculus is to develop a formalism to work with non-linear systems as if they were linear ones. The idea is to substitute the non-linear space M with the algebra $C^\infty(M)$. Points become *linear functionals*, vector fields *derivations* and diffeomorphisms are mapped to *automorphisms*. For proofs, and issues as topology used and continuity, check [12][Chapter 2]. We can rephrase the problem of solving an ODE on M as finding the solution to a *linear* ODE for operators on $C^\infty(M)$. To do so let us introduce the following dictionary:

$$\begin{aligned} q \in M &\iff \hat{q} \in (C^\infty(M))^*, & \hat{q}(f) &= f(q). \\ X \in \mathrm{Vec}(M) &\iff \hat{X} \in \mathrm{Der}(C^\infty(M)), & \hat{X}(f) &= df(X). \\ F \in \mathrm{Diff}(M) &\iff \hat{F} \in \mathrm{Aut}(C^\infty(M)), & \hat{F}(f) &= f \circ F. \end{aligned}$$

Let $F \in \mathrm{Diff}(M)$, it acts on $\mathrm{Vec}(M)$ as follows: $F_*(X)_{F(p)} = d_p F \, X_p$. This means that, as derivations, $\hat{F} \circ (F_* \hat{X}) = \hat{X} \circ \hat{F}$. This action is the so called *adjoint action* of $\mathrm{Aut}(C^\infty(M))$ on Der(C$^\infty$(M)) and we denote it by $\mathrm{Ad}\,\hat{F}^{-1}$. On $\mathrm{Der}(C^\infty(M))$ we have a natural structure of Lie algebra given by the commutator on vector fields. We can think of Der(C$^\infty$(M)) as the Lie algebra of $\mathrm{Aut}(C^\infty(M))$.

Any ODE $\dot{q} = V_t(q)$ can then be rephrased as an equation for its flow, seen as a linear operator on $C^\infty(M)$:

$$\frac{d}{dt}\hat{\Phi}_t = \hat{\Phi}_t \circ V_t, \quad \hat{\Phi}_0 = I.$$

The solution to this equation is unique and it is denoted by $\overrightarrow{\exp} \int_0^t V_\tau d\tau$, the right chronological exponential. It is possible to give an asymptotic expansion of the solution, the so called *Volterra series*. The Cauchy problem above in integral form reads:

$$\Phi_t = I + \int_0^t \Phi_\tau \circ V_\tau d\tau = I + \int_0^t V_\tau d\tau + \int_0^t \int_0^{t_1} \Phi_{t_1} \circ V_{t_1} \circ V_\tau d\tau dt_1.$$

Defining the symplex $\Delta_n(t) = \{t \geq t_1 \geq \cdots \geq t_n \geq 0\}$ and iterating we get, for $m \in \mathbb{N}$:

$$\Phi_t = I + \sum_{n=0}^{m-1} \int_{\Delta_n(t)} V_{t_n} \circ \cdots \circ V_\tau dt_n \dots d\tau + \int_{\Delta_m(t)} \Phi_{t_m} \circ V_{t_m} \circ \cdots \circ V_\tau dt_m \dots d\tau$$

We will use this together with the following formula for the variation of vector fields to differentiate a vector field depending on a parameter. Suppose you have two fields V_t and W_t, one can show (see [12][Section 2.7]), that the solution of $\dot{\Psi}_t = \Psi_t \circ (V_t + W_t)$, can be expressed in terms of $\overrightarrow{\exp} \int_0^t V_\tau d\tau$ as follows:

$$\overrightarrow{\exp} \int_0^t V_\tau + W_\tau d\tau = \overrightarrow{\exp} \int_0^t \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_r dr \right) W_\tau d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau d\tau.$$

Expanding the first term using Volterra series, we find that:

$$\overrightarrow{\exp} \int_0^t \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_r dr \right) W_\tau d\tau = I + \int_0^t \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_r dr \right) W_\tau d\tau$$
$$+ \int_0^t \int_0^\tau \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\theta V_r dr \right) W_\theta \circ \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_s ds \right) W_\tau d\theta d\tau + \dots$$

We can use this formula to obtain expression for the derivatives of a flow depending on parameter ϵ. Assume that $V_\tau(\epsilon)$ is a vector field depending on time and the parameter ϵ, for ϵ_0 fixed we can write $V_\tau(\epsilon) = V_\tau(\epsilon_0) + V_\tau(\epsilon) - V_\tau(\epsilon_0)$. Apply the formula above with $W_\tau = V_\tau(\epsilon) - V_\tau(\epsilon_0)$, to compute the value of the derivative one has to differentiate:

$$\overrightarrow{\exp} \int_0^t \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_r(\epsilon_0) dr \right) (V_\tau(\epsilon) - V_\tau(\epsilon_0)) d\tau \tag{B.4}$$

Now Volterra series comes in help. Since $W_\tau(\epsilon_0) = 0$, when we differentiate term by term the above series only a finite number of terms appear, in particular for the first derivative we have:

$$\partial_\epsilon \overrightarrow{\exp} \int_0^t V_\tau(\epsilon_0) d\tau = \int_0^t \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^\tau V_r(\epsilon_0) dr \right) \partial_\epsilon V_\tau(\epsilon_0) d\tau \circ \overrightarrow{\exp} \int_0^t V_\tau(\epsilon_0) d\tau \tag{B.5}$$

Set $P_\tau(\epsilon_0) = \overrightarrow{\exp} \int_0^\tau V_r(\epsilon_0) dr$ to simplify a bit the notation. For the second derivative, differentiating twice eq. (B.4) yields:

$$\int_0^t \mathrm{Ad}\, P_\tau(\epsilon_0) \partial_\epsilon^2 V_\tau(\epsilon_0) d\tau + 2 \int_0^t \int_0^\tau \mathrm{Ad}\, P_\theta(\epsilon_0) \partial_\epsilon V_r(\epsilon_0) \circ \mathrm{Ad}\, P_\tau(\epsilon_0) \partial_\epsilon V_\tau(\epsilon_0) d\theta d\tau \tag{B.6}$$

This formulas are the main tools we will use to get integral expression for the derivatives of the endpoint map defined in the previous sections.

B.2 Differentiation of the Endpoint map

Now we apply the formulas, especially eq. (B.5), of the previous section to compute the first and second differentials of the Endpoint map. Assume an extremal λ_t, either *normal* or *abnormal*, has been fixed. By point *iv*) of Theorem B.1 it solves the equation:

$$\dot{\lambda}_t = \vec{h}^t_{\tilde{u}}(\lambda_t)$$

We have to differentiate at the point $u = \tilde{u}(t)$, the optimal control. Take another control v, by the formula in eq. (B.5), we can write the Endpoint map at $\tilde{u} + \epsilon v$ as follow. We stress that this is an equality of linear functionals on $C^\infty(M)$.

$$E^t_{q_0}(\tilde{u} + \epsilon v) = q_0 \circ \overrightarrow{\exp} \int_0^t \mathrm{Ad}\left(\overrightarrow{\exp} \int_0^\tau V_r(\epsilon) dr\right) \partial_\epsilon V_\tau(\epsilon) d\tau \circ \overrightarrow{\exp} \int_0^t f_{\tilde{u}}\, d\tau.$$

Here the one parameter family of vector fields is given by:

$$V_t(\epsilon)(q) = f_{\tilde{u}(t)+\epsilon v(t)}(q) - f_{\tilde{u}}(q) \Rightarrow \begin{cases} \partial_\epsilon V_t(\epsilon)|_{\epsilon=0} = \partial_u f_{\tilde{u}(t)}\, v(t), \\ \partial^2_\epsilon V_t(\epsilon)|_{\epsilon=0} = \partial^2_u f_{\tilde{u}(t)}\, (v(t), v(t)). \end{cases}$$

Let us call the flow generate by $f_{\tilde{u}(t)}$, P_t. Clearly $P_t(q_0) = E^t_{q_0}(\tilde{u}) := q_t$ gives back the starting critical trajectory at time t. We have seen in the previous section that the action of $\mathrm{Ad}\, F$ on vector fields is the same as the push-forward of F^{-1}. In particular we have that:

$$q_0 \circ \mathrm{Ad}\left(\overrightarrow{\exp} \int_0^\tau V_r(\epsilon) dr\right) \partial_\epsilon V_\tau(\epsilon)|_{\epsilon=0} = q_0 \circ P_\tau \circ \partial_u f_{\tilde{u}(\tau)} v(\tau) \circ P_\tau^{-1}.$$

Now take $f \in C^\infty(M)$ and notice that we have to exchange the order in the former expression since when we differentiate $X \circ P(f) = X(f \circ P) = df\, dP\, X$. This implies that the differential of the endpoint map is:

$$d_{\tilde{u}} E^t_{q_0}\, v = (P_t)_* \int_0^t (P_\tau)^{-1}_* \partial_u f_{\tilde{u}}(q_\tau)\, v(\tau) d\tau. \tag{B.7}$$

For the second derivative we use eq. (B.6), substituting $V_t(\epsilon)$ and its derivatives and evaluating in zero yields:

$$\begin{aligned} d^2_{\tilde{u}} E^t_{q_0}(v, v) =& (P_t)_* \Bigg(\int_0^t (P_\tau)^{-1}_*\, \partial^2_u f_{\tilde{u}(\tau)}\, (v(\tau), v(\tau)) d\tau + \\ &+ 2 \int_0^t \int_0^\tau (P_\theta)^{-1}_* \partial_u f_{\tilde{u}}(q_\theta)\, v(\theta) \circ (P_\tau)^{-1}_* \partial_u f_{\tilde{u}}(q_\tau)\, v(\tau) d\tau\, dt \Bigg). \end{aligned} \tag{B.8}$$

Remark B.2 (Intrinsic Hessian). It is well known that the first differential of a map is a well defined linear map, independent on the coordinates chosen. This is no longer true for the second differential. For example, given any smooth function f on $\mathbb{R}^d$, there is always a change of coordinates that make it linear in a neighbourhood of a regular point. The second derivative of function is a well defined quadratic form on the kernel of the first differential, with values in the cokernel (i.e. the space $\mathrm{Codomain}(df)/\mathrm{Im}(df)$). Thus the second derivative will be the *family* of quadratic forms $\lambda_t d^2_{\tilde{u}} E^t_{q_0}$ for λ_t such that $\lambda_t d_{\tilde{u}} E^t_{q_0} \equiv 0$ defined on $\ker(d_{\tilde{u}} E^t_{q_0})$. (for further details see [12][Chapter 20])

The second term in eq. (B.8) can be rewritten in terms of *commutator* of vector fields. In fact if we split $[0, t] \times [0, t]$ in two simplexes, say $\Delta_< = \{(\tau, \theta) : \theta < \tau\}$ and $\Delta_> = \{(\tau, \theta) : \theta > \tau\}$, we have the following equalities for every $f \in C^\infty(M)$:

$$\begin{aligned} \left(\int_0^t X_\theta d\theta\right) \circ \left(\int_0^t Y_\tau d\tau\right)(f) &= \int_0^t \int_0^t X_\theta(Y_\tau(f)) d\theta d\tau \\ &= \int_{\Delta_<} X_\theta \circ Y_\tau(f) d\theta d\tau + \int_{\Delta_>} X_\theta \circ Y_\tau(f) d\theta d\tau \\ &= \int_{\Delta_<} (X_\theta \circ Y_\tau + X_\tau \circ Y_\theta)(f) d\theta d\tau. \end{aligned}$$

In particular this implies the following:

$$2\int_0^t\int_0^\tau X_\theta \circ Y_\tau = \int_0^t X_\theta d\theta \circ \int_0^t Y_\tau d\tau + \int_0^t\int_0^\tau [X_\tau, Y_\theta] d\tau d\theta.$$

The second derivative is intrinsically defined on the kernel of the first derivative, with values in its co-kernel (see remark B.2). We are thus interested in the family of quadratic forms $\eta d^2_{\tilde{u}} E^t_{q_0}$, where $\eta d_{\tilde{u}} E^t_{q_0} \equiv 0$. In particular, if λ_t is our extremal eq. (B.8), the second differential can be written as:

$$\begin{aligned}\lambda_t \, d^2_{\tilde{u}} E^t_{q_0}(v,v) =& \lambda_t (P_t)_* \Bigg(\int_0^t (P_\tau)_*^{-1} \, \partial_u^2 f_{\tilde{u}(\tau)} \, (v(\tau), v(\tau)) d\tau + \\ &+ \int_0^t \int_0^\tau [(P_\theta)_*^{-1} \partial_u f_{\tilde{u}}(q_\theta) \, v(\theta), (P_\tau)_*^{-1} \partial_u f_{\tilde{u}}(q_\tau) \, v(\tau)] d\tau \, dt \Bigg).\end{aligned} \tag{B.9}$$

By point $iv)$ of Theorem B.1 follows that $\lambda_t (P_t)_* = \lambda_0$. Moreover we can rewrite the formula above in symplectic terms, using the functions $h^t_u(\lambda) = \langle \lambda, f^t_u \rangle$ appearing in the statement of PMP. For a vector field Z, possibly depending on extra parameters u and on time t, define the function $h_Z(\lambda) := \langle \lambda, Z \rangle$ on T^*M. One can check (see again [12]) the following identities:

$$\begin{aligned}\partial_u^2 h_{Z_u}(\lambda) &= \langle \lambda, \partial_u^2 Z_u \rangle, \\ h_{[X_t, Y_t]}(\lambda) &= \langle \lambda, [X_t, Y_t] \rangle = \sigma_\lambda(\vec{h}_X, \vec{h}_Y).\end{aligned} \tag{B.10}$$

The only issue to address now is to interpret eq. (B.9) in an intrinsic way. Notice that the objects appearing as integrands always live in the tangent space to the initial point λ_0. This is why we are allowed to speak about *integrals*. We use now the solution of $\dot{\lambda} = \vec{h}^t_{\tilde{u}}(\lambda)$ to perform a time dependent change of coordinates that backtracks the extremal λ_t to λ_0 and rewrite the integrands in term of this flow, using the just mentioned identities. Let us call this flow $\tilde{\Phi}_t$. This procedure is in some equivalent to fixing a connection and defining a *parallel transport* operator which trivializes the bundle $\lambda_t^* T(T^*M)$. Define the Hamiltonian:

$$b^t_u(\lambda) = (h^t_u - h^t_{\tilde{u}}) \circ \tilde{\Phi}_t(\lambda)$$

Combining with the identities written in eq. (B.10) we have the following equalities (see [12][Chapter 20.3]):

$$\begin{aligned}\langle \lambda_0, [(P_\theta)_*^{-1} \partial_u f_{\tilde{u}}(q_\theta) \, v(\theta), (P_\tau)_*^{-1} \partial_u f_{\tilde{u}}(q_\tau) \, v(\tau)] \rangle &= \sigma_{\lambda_0}(\partial_u \vec{b}^\theta_u, \partial_u \vec{b}^\tau_u) \\ \langle \lambda_0, (P_\tau)_*^{-1} \partial_u^2 f_{\tilde{u}}(v(t), v(t)) \rangle &= \partial_u^2 b^t_u(\lambda_0)(v(t), v(t)),\end{aligned}$$

By the discussion in the previous sections (see eq. (B.3)), we can formulate optimal control problems in terms of the endpoint map of an auxiliary system. Thus to get the second variation of an extremal for an optimal control problem we have to consider the following fields on $M \times \mathbb{R}$:

$$\hat{f}^t_u(q,x) = \begin{pmatrix} \varphi(t,u,q) \\ f^t_u(q) \end{pmatrix}$$

It turns out that this is equivalent to using the Hamiltonians $h^t_u(\lambda) = \langle \lambda, f^t_u \rangle + \nu \varphi(t,u,q)$ appearing in the statement of Theorem B.1 on T^*M. Through out the thesis we will use the following notation for the matrices appearing in the expression of the second derivative. We summarize it all here for further references:

$$\begin{aligned}
&\textbf{Hamiltonians of PMP: } h_u^t(\lambda) = \langle \lambda, f_u^t\rangle + \nu\varphi(t,u,q), \quad h_{\tilde u}^t(\lambda) = h_{\tilde u(t)}^t(\lambda),\\
&\textbf{Re-parametrization flow (or \textit{parallel transport}): } \frac{d}{dt}\tilde\Phi_t = \tilde\Phi_t \circ \vec{h}_{\tilde u}^t,\\
&\textbf{Re-parametrized Hamiltonian: } b_u^t(\lambda) = (h_u^t - h_{\tilde u}^t)\circ\tilde\Phi_t(\lambda)\\
&\textbf{Second Variation's coefficients: } Z_t := \partial_u \vec{b}_u^t(\lambda)|_{\lambda=\lambda_0}, \quad H_t := \partial_u^2 b_u^t(\lambda)|_{\lambda=\lambda_0}.
\end{aligned} \tag{B.11}$$

We collect here formulas for the first two derivatives of the endpoint map.

Proposition B.2 (Derivatives of the Enpoint map). *Suppose that $(\tilde u_t, \lambda_t)$ is a couple of optimal control and extremal for the optimal control problem in eq.* (B.2) *with Dirichlet boundary conditions. With the notation introduced above and summarized in eq.* (B.11) *the first differential of the Endpoint map is given by the following map:*

$$d_{\tilde u}E_{q_0}^t(v) = \pi_*(\tilde\Phi_t)_*\int_0^t Z_\tau v_\tau d\tau.$$

The second differential is given by the quadratic mapping in eq. (B.8)*. Moreover, for $v \in \ker d_{\tilde u}E_{q_0}^t$, the projection along the extremal λ_t satisfies:*

$$\lambda_t d_{\tilde u}^2 E_{q_0}^t(v,v) = \int_0^t \left(\langle H_\tau v_\tau, v_\tau\rangle + \int_0^\tau \sigma_{\lambda_0}(Z_\theta v_\theta, Z_\tau v_\tau)d\theta\right)d\tau.$$

B.3 Classes of extremals

In this section we collect some definitions we will use later.

Definition B.2. An extremal λ_t with optimal control $\tilde u(t)$ is called:

i) *regular* if $H_t(v,v) < -\alpha\|v\|^2$ for $\alpha > 0$ and $v \in \mathbb{R}^k$,

ii) *strictly normal* if λ_t is a normal extremal and the differential of the endpoint map is surjective

iii) *totally singular* if $H_t \equiv 0$.

The condition given in $i)$ is often referred to as *Legendre strong condition* and is a sufficient condition for the finiteness of the index of $\lambda_t d_{\tilde u}^2 E_{q_0}^t$.

Definition B.3. We say that an extremal is of corank $m \geq 1$ if the space $T_{q_t}M/\mathrm{Im}(d_{\tilde u}E_{q_0}^t)$ has dimension m.

In particular strictly normal extremal have always corank 1. In this case the second derivative is actually a quadratic form on a subspace of $L^\infty([0,t],\mathbb{R}^k)$ of codimension $\dim(M)$.

Bibliography

[1] R. Adami. Ground states for NLS on graphs: a subtle interplay of metric and topology. *Math. Model. Nat. Phenom.*, 11(2):20–35, 2016.

[2] Riccardo Adami, Claudio Cacciapuoti, Domenico Finco, and Diego Noja. Variational properties and orbital stability of standing waves for NLS equation on a star graph. *J. Differential Equations*, 257(10):3738–3777, 2014.

[3] Riccardo Adami, Enrico Serra, and Paolo Tilli. NLS ground states on graphs. *Calc. Var. Partial Differential Equations*, 54(1):743–761, 2015.

[4] Riccardo Adami, Enrico Serra, and Paolo Tilli. Threshold phenomena and existence results for NLS ground states on metric graphs. *J. Funct. Anal.*, 271(1):201–223, 2016.

[5] A. Agrachev, G. Stefani, and P. Zezza. An invariant second variation in optimal control. *Internat. J. Control*, 71(5):689–715, 1998.

[6] A. A. Agrachëv. Quadratic mappings in geometric control theory. In *Problems in geometry, Vol. 20 (Russian)*, Itogi Nauki i Tekhniki, pages 111–205. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1988. Translated in J. Soviet Math. **51** (1990), no. 6, 2667–2734.

[7] A. A. Agrachev. Spectrum of the second variation. *Tr. Mat. Inst. Steklova*, 304(Optimal noe Upravlenie i Differentsial nye Uravneniya):32–48, 2019.

[8] Andrei Agrachev, Davide Barilari, and Ugo Boscain. *A comprehensive introduction to sub-Riemannian geometry*, volume 181 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2020. From the Hamiltonian viewpoint, With an appendix by Igor Zelenko.

[9] Andrei Agrachev and Ivan Beschastnyi. Jacobi fields in optimal control: one-dimensional variations. *J. Dyn. Control Syst.*, 26(4):685–732, 2020.

[10] Andrei Agrachev and Ivan Beschastnyi. Jacobi fields in optimal control: Morse and Maslov indices. *Nonlinear Anal.*, 214:Paper No. 112608, 47, 2022.

[11] Andrei A. Agrachev, A. Stephen Morse, Eduardo D. Sontag, Héctor J. Sussmann, and Vadim I. Utkin. *Nonlinear and Optimal Control Theory: Lectures Given at the C.I.M.E. Summer School Held in Cetraro, Italy, June 19-29, 2004*. Lectures notes in mathematics. Springer Nature., 2008.

[12] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[13] Andrei A. Agrachev, Gianna Stefani, and Pierluigi Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optim.*, 41(4):991–1014, 2002.

[14] Andrey A. Agrachev and Ivan Yu. Beschastnyi. Symplectic geometry of constrained optimization. *Regul. Chaotic Dyn.*, 22(6):750–770, 2017.

[15] Baranzini S. Agrachev A. and I. Beschastnyi. Index theorems for graph-parametrized optimal control problems. *Nonlinearity*, Submitted.

[16] Vladimir I Arnol. The sturm theorem and symplectic geometry. 2004.

[17] Stefano Baranzini. Functional determinants for the second variation. *submitted to Journal of Differential Equations*, 2022.

[18] Stefano Baranzini. Operators arising as second variation of optimal control problems and their spectral asymptotics. *Journal of Dynamical and Control Systems*, 07 2022.

[19] Yu. M. Baryshnikov. Indices for extremal embeddings of 1-complexes. In *Theory of singularities and its applications*, volume 1 of *Adv. Soviet Math.*, pages 137–144. Amer. Math. Soc., Providence, RI, 1990.

[20] S. V. Bolotin and D. V. Treshchëv. Hill's formula. *Uspekhi Mat. Nauk*, 65(2(392)):3–70, 2010.

[21] William Borrelli, Raffaele Carlone, and Lorenzo Tentarelli. An overview on the standing waves of nonlinear schrödinger and dirac equations on metric graphs with localized nonlinearity. *Symmetry*, 11(2), 2019.

[22] R. Bott. On the iteration of closed geodesics and the Sturm intersection theory. *Comm. Pure Appl. Math.*, 9:171–206, 1956.

[23] Claudio Cacciapuoti, Domenico Finco, and Diego Noja. Ground state and orbital stability for the NLS equation on a general starlike graph with potentials. *Nonlinearity*, 30(8):3271–3303, 2017.

[24] Charles K. Chui. Concerning rates of convergence of Riemann sums. *J. Approximation Theory*, 4:279–287, 1971.

[25] Charles C. Conley and Eduard Zehnder. Morse-type index theory for flows and periodic solutions for hamiltonian equations. *Communications on Pure and Applied Mathematics*, 37:207–253, 1984.

[26] J.B. Conway. *Functions of One Complex Variable I*. Functions of One Complex Variable. Springer, 1978.

[27] Maurice de Gosson. *Symplectic geometry and quantum mechanics*, volume 166 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2006. Advances in Partial Differential Equations (Basel).

[28] Simone Dovetta. Mass-constrained ground states of the stationary NLSE on periodic metric graphs. *NoDEA Nonlinear Differential Equations Appl.*, 26(5):Paper No. 30, 30, 2019.

[29] Simone Dovetta, Enrico Serra, and Paolo Tilli. Nls ground states on metric trees: existence results and open questions. *J. London Math. Society*, (3), 05 2019.

[30] J. J. Duistermaat. On the Morse index in variational calculus. *Advances in Math.*, 21(2):173–195, 1976.

[31] E. Elizalde, S. D. Odintsov, A. Romeo, A. A. Bytsenko, and S. Zerbini. *Zeta regularization techniques with applications*. World Scientific Publishing, Singapore, 1994.

[32] R. Forman. Functional determinants and geometry . *Inventiones mathematicae*, 88:447–494, 1987.

[33] Leonid Friedlander. Determinant of the schrödinger operator on a metric graph. *Contemporary Mathematics*, 01 2006.

[34] Luiz Hartmann and Matthias Lesch. Zeta and fredholm determinants of self-adjoint operators. *Journal of Functional Analysis*, 283(1):109491, 2022.

[35] G. W. Hill. On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon. *Acta Mathematica*, 8(none):1 – 36, 1900.

[36] Xijun Hu, Yuwei Ou, and Penghui Wang. Trace formula for linear Hamiltonian systems with its applications to elliptic Lagrangian solutions. *Arch. Ration. Mech. Anal.*, 216(1):313–357, 2015.

[37] Xijun Hu, Yuwei Ou, and Penghui Wang. Hill-type formula for hamiltonian system with lagrangian boundary conditions. *Journal of Differential Equations*, 267(4):2416–2447, 2019.

[38] Xijun Hu and Penghui Wang. Conditional fredholm determinant for the s-periodic orbits in hamiltonian systems. *Journal of Functional Analysis*, 261(11):3247–3278, 2011.

[39] Xijun Hu and Penghui Wang. Conditional fredholm determinant for the s-periodic orbits in hamiltonian systems. *Journal of Functional Analysis - J FUNCT ANAL*, 261:3247–3278, 12 2011.

[40] Xijun Hu and Penghui Wang. Hill-type formula and krein-type trace formula for s-periodic solutions in odes. *Discrete and Continuous Dynamical Systems*, 36(2):763–784, 2016.

[41] Xijun Hu, Li Wu, and Ran Yang. Morse index theorem of lagrangian systems and stability of brake orbit. *Journal of Dynamics and Differential Equations*, 32, 03 2020.

[42] Xijun Hu, Li Wu, and Ran Yang. Morse index theorem of lagrangian systems and stability of brake orbit. *Journal of Dynamics and Differential Equations*, 32, 03 2020.

[43] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. SpringerBriefs in Mathematics. Springer, Cham, 2014.

[44] Tosio Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[45] K. Kirsten and Alan McKane. Functional determinants for general sturm-liouville problems. *Journal of Physics A General Physics*, 37, 04 2004.

[46] Klaus Kirsten and Alan J. McKane. Functional determinants by contour integration methods. *Annals of Physics*, 308:502–527, 2003.

[47] MG Krein. The basic propositions of the theory of λ-zones of stability of a canonical system of linear differential equations with periodic coefficients. In *Topics in differential and integral equations and operator theory*, pages 1–105. Springer, 1983.

[48] M. G. Kreĭn. On criteria of stable boundedness of solutions of periodic canonical systems. *Prikl. Mat. Meh.*, 19:641–680, 1955.

[49] Yiming Long. *Index Theory for Symplectic Paths with Applications.* 01 2002.

[50] M. Morse. *The Calculus of Variations in the Large.* Number v. 18 in American Mathematical Society. American Mathematical Society, 1934.

[51] Daniel Offin. Hyperbolic minimizing geodesics. *Transactions of the American Mathematical Society*, 352(7):3323–3338, 2000.

[52] H. Poincaré. Les méthodes nouvelles de la mécanique céleste. Tome I. Solutions périodiques. Non-existence des intégrales uniformes. Solutions asymptotiques. Paris: Gauthier-Villars et Fils. 385 p. gr. 8° (1892)., 1892.

[53] Joel Robbin and Dietmar Salamon. The maslov index for paths. *Topology*, 32(4):827–844, 1993.

[54] Walter Rudin. *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[55] B. Simon and S.P.G.N.J. Hitchin. *Trace Ideals and Their Applications.* Lecture note series / London Mathematical Society. Cambridge University Press, 1979.

[56] D. V. Treshchev. Connection of the Morse index of a closed geodesic line with its stability. *Tr. Semin. Vektorn. Tenzorn. Anal.*, 23:175–189, 1988.

[57] Antonio Ureña. All periodic minimizers are unstable. *Archiv Der Mathematik - ARCH MATH*, 91:63–75, 07 2008.

[58] Antonio Ureña. The spectrum of reversible minimizers. *Regular and Chaotic Dynamics*, 23:248–256, 05 2018.

[59] Vladimir A. Yakubovich. On m. g. krein's work in the theory of linear periodic hamiltonian systems. *Ukrainian Mathematical Journal*, 46:133–148, 1994.