Journal of Dynamical and Control Systems (2023) 29:659–689
https://doi.org/10.1007/s10883-022-09605-1

Operators Arising as Second Variation of Optimal Control Problems and Their Spectral Asymptotics

Stefano Baranzini[1]

Received: 14 September 2021 / Revised: 26 April 2022 / Accepted: 4 June 2022 /
Published online: 8 July 2022

Abstract

We compute the asymptotic for the eigenvalues of a particular class of compact operators deeply linked with the second variation of optimal control problems. We characterize this family in terms of a set of finite dimensional data and we apply this results to a particular class of singular extremal to get a nice description of the spectrum of the second variation.

Keywords Second variation · Optimal control · Weyl law · Compact operator

Mathematics Subject Classification (2010) 47A11 · 47A53 · 47G10 · 49K99

1 Introduction

The main focus of this paper is the study of a particular class of compact operators K on the Hilbert space $L^2([0,1],\mathbb{R}^k)$ with the standard Hilbert structure. They are characterized by the following properties:

- there exists a finite 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. :
$$\langle u, Kv\rangle = \langle Ku, v\rangle \quad \forall u, v \in \mathcal{V}, \tag{1}$$
- K is an Hilbert-Schmidt operator with an integral kernel of a particular form, namely:
$$K(v)(t) = \int_0^t V(t,\tau)v(\tau)d\tau, \quad v \in L^2([0,1],\mathbb{R}^k), \tag{2}$$
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.

The main results of this paper are a fairly general study of the asymptotic distribution of the eigenvalues of K when restricted to any subspace $\mathcal{V}$ which satisfies (1)

✉ Stefano Baranzini
sbaranzi@sissa.it

[1] SISSA, Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea, 265 - 34136, Trieste, Italy

(Theorem 1) and a characterization result for operators satisfying the two properties stated above (Theorem 2).

The first result is proved in Section 3. We first restrict ourself to operators $\tilde{K}$ of the form:

$$\tilde{K}(v)(t) = -\int_0^t \sigma(Z_\tau v_\tau, Z_t \cdot)d\tau. \tag{3}$$

Here, Z_t is an analytic in t, $2n \times k$ matrix and σ the standard symplectic form on $\mathbb{R}^{2n}$ (see Remark 1). A similar asymptotic formula was proved in [3, Theorem 1], it was shown that if we consider $\{\lambda_n(\tilde{K})\}_{n\in\mathbb{Z}}$ the decreasing (resp. increasing) arrangement of positive (resp. negative) eigenvalues of $\tilde{K}$ we have either:

$$\lambda_n(\tilde{K}) = \frac{\xi}{\pi n} + O(n^{-5/3}) \quad \text{or} \quad \lambda_n(\tilde{K}) = O(n^{-2}), \tag{4}$$

for $n \in \mathbb{Z}$ sufficiently large and for some $\xi > 0$. The number ξ is called *capacity* and depends only on the matrix Z_t in the definition of $\tilde{K}$.

If $\xi = 0$, we go further with the expansion in (4). We single out the term giving the principal contribution to the asymptotic representing the quadratic form associated to $\tilde{K}$ as:

$$Q(v) = \langle v, \tilde{K}v\rangle = -\int_0^1\int_0^t \sigma(Z_\tau v_\tau, Z_t u_t)d\tau dt = \sum_{i=1}^{k-1} Q_i(v) + R_k(v).$$

The result mentioned above corresponds to the case $Q_1 \neq 0$; in Theorem, 1 we give the asymptotic for the general case.

From the point of view of geometric control theory, Theorem 1 can be seen as an asymptotic analysis of the spectrum of the second variation for particular classes of singular extremals and a quantitative version of some necessary optimality conditions.

Precise definitions will be given in Section 5. Standard references on the second variation are [7, Chapter 20] and [1]. For now, it is enough to know that the second variation Q of an optimal control problem on a manifold M is a linear operator on $L^2([0,1], \mathbb{R}^k)$ of the following form:

$$\langle Qv, u\rangle = -\int_0^1 \langle H_t v_t, u_t\rangle - \int_0^1\int_0^t \sigma(Z_\tau v_\tau, Z_t u_t)d\tau dt, \tag{5}$$

where H_t is a symmetric $k \times k$ matrix, σ is the standard symplectic form on $T_\eta T^*M$ and $Z_t : \mathbb{R}^k \to T_\eta(T^*M)$ is a linear map with values in the tangent space to a fixed point $\eta \in T^*M$.

For totally singular extremal, the matrix H_t appearing in (5) is identically zero and the second variation reduces to an operator of the same form as in (3).

In Section 4, we prove Theorem 2. We first show that any K satisfying (1) and (2) it is completely determined by its (*finite rank*) skew-symmetric part $\mathcal{A}$ and can always be represented as in (3). Then we relate the *capacity* of K to the spectrum of $\mathcal{A}$.

In Section 5, we recall some basic notions from control theory and we reformulate Theorem 2 in a more control theoretic fashion, and use it to characterize the operators coming form the *second variation* of an optimal control problem. Moreover, we give a geometric interpretation of the capacity ξ appearing in (4) in terms of the Hessian of the maximized Hamiltonian coming from Pontryagin Maximum Principle.

2 Overview of the Main Results

We begin this section recalling some general facts about the spectrum of compact operators, then we fix some notation and give a precise statement of the main results. Given a compact self-adjoint operator K on an Hilbert space $\mathcal{H}$, we can define a quadratic form setting $Q(v) = \langle v, K(v)\rangle$. The eigenvalues of Q are by definition those of K and we will denote $\Sigma_\pm(Q)$ the positive and negative parts of the spectrum of Q.

By the standard spectral theory of compact operators (see [12]), 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 Q, $\Sigma_+(Q)$ and $\lambda \in \Sigma_+(Q)$. 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 $\lambda \in \Sigma_+(Q)$.

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

Definition 1 Let Q be a quadratic form Q on a Hilbert space $\mathcal{H}$ and $j \in \mathbb{N}$

- if j is odd, Q has $j-$capacity $\xi > 0$ with reminder of order $\nu > 0$ if $\Sigma_+(Q)$ and $\Sigma_-(Q)$ are both infinite and:
$$\lambda_n = \frac{\xi}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n \to \pm\infty,$$
- if j is even, Q has $j-$capacity (ξ_+, ξ_-) of order $\nu > 0$ if both $\Sigma_+(Q)$ and $\Sigma_-(Q)$ are infinite and:
$$\lambda_n = \tfrac{\xi_+}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n \to +\infty,$$
$$\lambda_n = \tfrac{\xi_-}{(\pi n)^j} + O(n^{-\nu-j}) \quad \text{as} \quad n \to -\infty,$$
where $\xi_\pm \geq 0$ or if at least one between $\Sigma_+(Q)$ and $\Sigma_-(Q)$ is infinite and the relative monotone arrangement satisfies the corresponding asymptotic relation;
- if the spectrum is finite or $\lambda_n = O(n^{-\nu})$ as $n \to \pm\infty$ for any $\nu > 0$, we say that 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 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 illustrate here some of the properties of the $j-$capacity. The proofs are given in Section 3, Proposition 3. Without loss of generality we state the properties for the positive part of the spectrum, analogue results hold for the negative one.

- *(Homogeneity)* if Q_1 and 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 aQ_1 has $j-$capacity

$a\xi_1$ and the sum $Q_1 \oplus Q_2$ on $\mathcal{H}_1 \oplus \mathcal{H}_2$ has $j-$capacity $\left(\sqrt[j]{\xi_1} + \sqrt[j]{\xi_2}\right)^j$ both with remainder ν.

- *(Independence of restriction)* If $\mathcal{V} \subseteq \mathcal{H}$ is a subspace of finite codimension then Q has $j-$capacity ξ with remainder ν if and only if its restriction to $\mathcal{V}$ has $j-$capacity ξ with remainder ν.
- *(Additivity)* if Q_1 has $j-$capacity ξ with remainder ν and Q_2 has 0 $j-$capacity with remainder of the same order ν, then their sum $Q_1 + Q_2$ has the same capacity with remainder $\nu' = \frac{(j+\nu)(j+1)}{j+\nu+1}$

The remaining part of this section will be dealing with quadratic forms Q coming from operators of the form given in (3). Suppose that Z_t is a $2n \times k$ matrix which depends piecewise analytically on the parameter $t \in [0, 1]$ and define the following $2n \times 2n$ skew-symmetric matrix:

$$J = \begin{pmatrix} 0 & -Id_n \\ Id_n & 0 \end{pmatrix}. \tag{6}$$

As Q consider the following quadratic form on $L^2([0, 1], \mathbb{R}^k)$:

$$Q(v) = \langle v, K(v)\rangle = \int_0^1 \int_0^t \langle Z_t v(t), J Z_\tau v(\tau)\rangle d\tau dt. \tag{7}$$

Remark 1 The operator K and the bilinear form $Q(u, v) = \langle u, K(v)\rangle$ are not symmetric. However, the operator:

$$K(v) = \int_0^t Z_t^* J Z_\tau v(\tau) d\tau,$$

satisfies (1) and becomes symmetric on a finite codimension subspace $\mathcal{V}$. It is enough to require that the integral $\int_0^1 Z_t v(t)dt$ lies in a Lagrangian subspace of $(\mathbb{R}^{2n}, \sigma)$ for any $v \in \mathcal{V}$. For instance, if we consider the fibre (or *vertical* subspace), i.e. the following:

$$\Pi = \{(p, 0) : p \in \mathbb{R}^n\} \subset \mathbb{R}^{2n}. \tag{8}$$

Here, σ denotes the standard symplectic form on $\mathbb{R}^{2n}$ defined as $\sigma(x, x') = \langle Jx, x'\rangle$.

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:

$$A_j(t) = \begin{cases} \left(Z_t^{(k)}\right)^* J Z_t^{(k)} & \text{if } j = 2k-1 \\ \left(Z_t^{(k-1)}\right)^* J Z_t^{(k)} & \text{if } j = 2k \end{cases} \tag{9}$$

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 first main result of the section.

Theorem 1 *Let Q be the quadratic form in (7). 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*

- *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:

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

- *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:

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

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

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

The proof of this result is given in Section 3. The next theorem gives a characterization of the operators satisfying (1) and (2) and a geometric interpretation of the $1-$capacity. Before going to the statement let us introduce the following notation. Let $\mathcal{A}$ denote the skew-symmetric part of K:

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

Let Σ be the spectrum of $\mathcal{A}$ and $\mathrm{Im}(\mathcal{A})$, the image of $\mathcal{A}$.

Theorem 2 *Let be K an operator satisfying (1) and (2). 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:

$$K(v)(t) = \int_0^t Z_t^* \mathcal{A}_0 Z_\tau v(\tau) d\tau. \tag{10}$$

Let Σ be the spectrum of $\mathcal{A}$, if the matrix Z_t can be chosen to be piecewise analytic the $1-$capacity of K can be bound by

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

3 Proof of Theorem 1

Before going to the proof of Theorem 1 we still need some auxiliary results. We start with Lemma 1 to single out the main contributions to the asymptotic of the eigenvalues of Q (the quadratic form defined in (7)). The first non-zero term of the decomposition we give will determine the rate of decaying of the eigenvalues (see Proposition 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).

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:

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

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

$$\langle \int_0^1 Z_t v(t) dt, J \int_0^1 \dot{Z}_t v_1(t) dt \rangle = \langle \int_0^1 Z_t v(t) dt, J Z_1 v_1(1)\rangle - \langle \int_0^1 Z_t v(t) dt, J \int_0^1 Z_t v(t) dt \rangle$$

and we can write Q as a sum of three terms

$$Q(v) = Q_1(v) + Q_2(v) + R_1(v)$$

In analogy, we can make the following definitions:

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

Here, the matrices $A_j(t)$ are exactly those defined in (9).

Lemma 1 *For every $j \in \mathbb{N}$, on the subspace V_j, the form Q can be represented as*

$$Q(v) = \sum_{k=1}^{2j} Q_k(v) + R_j(v) \tag{11}$$

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 $Q(v)$ 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. □

Now, we would like to compute explicitly the spectrum of the 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 Q_4 the equation determining it cannot be solved explicitly.

We will derive the Euler-Lagrange equation for 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 Q_j. If $j = 2k$ integration by parts yields:

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

Notice that the boundary terms vanish identically if we impose the vanishing of v_j for $1 \le j \le k$ 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 $1-$dimensional systems. Imposing periodic boundary conditions, we are thus left with the following boundary value problem:

$$w^{(2k)}(t) = \frac{(-1)^k \mu}{\lambda} w(t) \quad w^{(j)}(0) = w^{(j)}(1) \text{ for } 0 \le j \le 2k-1 \tag{12}$$

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

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

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 Q_{2k-1} as a direct sum of two-dimensional forms. Imposing periodic boundary conditions, we end up with the following boundary value problems:

$$\begin{cases} w_1^{(2k-1)}(t) &= -\frac{(-1)^{(k-1)}\mu}{\lambda} w_2 \\ w_2^{(2k-1)}(t) &= \frac{(-1)^{(k-1)}\mu}{\lambda} w_1 \end{cases} \qquad \begin{cases} w_1^{(j)}(0) = w_1^{(j)}(1), \\ w_2^{(j)}(0) = w_2^{(j)}(1) \end{cases} \text{for } 0 \le j \le 2k-2. \tag{13}$$

Lemma 2 *The boundary value problem in (12) 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 (12) 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 (13) 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:

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

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:

$$w^{(2k)}(t) = a^{2k} w(t), \qquad w^{(j)}(0) = w^{(j)}(1), \;\; 0 \le j < 2k. \tag{14}$$

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 $\{e^{\eta t}, e^{\bar{\eta} t}, e^{-\eta t}, e^{-\bar{\eta} t}\}$ 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 (14)) 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 $E = \{\eta : \Re(\eta) \geq 0, \Im(\eta) \geq 0, \eta^{2k} = 1\}$.

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 $\bigoplus_{\eta\neq i} U_\eta$. 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 arrange the solution in a decreasing order, 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:

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

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 arrange them in order, 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 (13) 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 arrange them in order, 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 1 *Let* $\mu > 0$ *and* $s \in (0, +\infty)$*, denote by* η_s *the number of solutions of* (12) *with* λ *greater than* s *and similarly denote by* ω_s *be the number of solutions with* λ *bigger than* s *of:*

$$w^{(2k)}(t) = \frac{(-1)^k\mu}{\lambda}w(t), \quad w^{(j)}(0) = w^{(j)}(1) = 0, \quad k \leq j \leq 2k-1 \tag{15}$$

Then, $|\omega_s - \eta_s| \leq 2k$*. The same conclusion holds for* (13).

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 [2, 5, 6]. 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 [2], the following inequality is proved:

$$|\mathrm{Mi}_{L_0}(\gamma) - \mathrm{Mi}_{L_1}(\gamma)| \leq \frac{\dim(\Sigma)}{2} \tag{16}$$

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 $\dot{\Phi}^t_\lambda = \hat{A}_\lambda \Phi^t_\lambda$ at time $t = 1$. Consider its graph:

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

Once we prove that γ is monotone, it is straightforward to check that $\mathrm{Mi}_{B\times B}(\gamma|_{[s,+\infty)})$ counts the number of solutions to boundary value problem given in (15) for $\lambda \geq s$ and similarly $\mathrm{Mi}_{\Gamma(I)}(\gamma|_{[s,+\infty)})$ counts the solutions of (12) 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:

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

where we used the facts that $\partial_\lambda \Phi^0_\lambda = \partial_\lambda Id = 0$ 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, it 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 (16). □

We are finally ready to compute the asymptotic for Q_j when the matrix A_j is constant. The next Proposition translates the estimate on the counting functions η_s and ω_s defined in Proposition 1 to an estimate for the eigenvalues.

Proposition 2 *Let Q_j be any of the forms appearing in* (11).

- *Suppose $j = 2k$ and $Q_{2k}(v) = \int_0^1 \langle A_{2k} v_k, v_k \rangle dt$ 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, 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 $r \geq mk$

$$\frac{\xi_+}{\pi^j(r-2mk-p(r))^j} \geq \lambda_r \geq \frac{\xi_+}{\pi^j(r+2mk+p(r))^j} \tag{17}$$

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

- *Suppose $j = 2k+1$ and $Q_{2k+1}(v) = \int_0^1 \langle A_{2k+1}v_{k-1}, v_k\rangle dt$ 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, 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 $|r| \geq mk$

$$\frac{\xi}{\pi^j(r-2mk-p(r))^j} \geq \lambda_r \geq \frac{\xi}{\pi^j(r+2mk+p(r))^j}. \tag{18}$$

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:

$$g(s) \geq \#\{\lambda \text{ solutions of (15) } : \lambda \geq s\} \geq f(s),$$

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 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:

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

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 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 $Q = \oplus_{i=1}^m Q_i$, λ is an eigenvalue of Q if and only if it is an eigenvalue of Q_i for some i. We proceed as we did before. Suppose that Q_a is $1-$dimensional and $Q_a(v) = \int_0^1 \mu_a |v_k(t)|^2 dt$. Let us compute η_s in the point $s_0 = \left(\sum_{i=1}^m \sqrt[j]{\mu_i}\right)^j / (2\pi l)^j$:

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

Set for simplicity $c_a = \frac{\sqrt[j]{\mu_a}}{\left(\sum_{i=1}^m \sqrt[j]{\mu_i}\right)}$, 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:

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

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:

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

It is straightforward to compute the bounds in (17) and (18) observing again $\lambda_{2l} \geq \lambda_{2l+1} \geq \lambda_{2l+2}$. □

Remark 3 The shift m appearing in (17) and (18) 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 $L^2([0,1], \mathbb{R}^m)$ and the estimates in (17) and (18) 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 below.

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

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

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

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

2. *Suppose that U is a subspace of codimension $d < \infty$ then*

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

as $n \to +\infty$.

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

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

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

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

Proof The asymptotic relation can be written in terms of a counting function. Take the jth root of the eigenvalues of Q_i, then it holds that

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

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

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

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

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

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 ith and the jth eigenvalues of the summands. Write, as in [3], Q' as $Q+\hat{Q}$ and Q as $Q'+(-\hat{Q})$. 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:

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

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 Q_j, independently of any analyticity assumption of the kernel.

Proposition 4 *Take Q_j as in the decomposition of lemma Eq. 1. Then, the eigenvalues of Q_j satisfy:*

$$\lambda_n(Q_j) = O\left(\frac{1}{n^j}\right) \quad \text{as } n \to \pm\infty$$

Moreover, for any $k \in \mathbb{N}$ and for any $0 \leq s \leq k$, the forms Q_{2k+1} and Q_{2k} have the same first term asymptotic as the forms:

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

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

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

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

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

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

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 nth 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.

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

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:

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

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. □

Now, we prove the main result of this section:

Proof of Theorem 1 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

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

Since the matrix A_t is analytic, we can diagonalize it piecewise analytically in t (see [11]). 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 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:

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

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 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:

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

Now, by Proposition 3 (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 $v_k^{t_i}(t) = \int_{t_i}^t \dots \int_{t_i}^{\tau_1} v(\tau)d\tau \dots d\tau_{k-1}$. It follows immediately that on V_Π:

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

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 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 (17) and (18) 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, we have the following estimate on the eigenvalues on V_Π, for $n \geq 2k|\Pi|$:

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

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

$$\lambda_n\left(Q_{2k}\big|_{V_\Pi}\right) \leq \lambda_n\left(Q_{2k}\right) \leq \lambda_{n-k|\Pi|}\left(Q_{2k}\big|_{V_\Pi}\right)$$

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

$$\left(\frac{\sum_i (a_i^-)^{\frac{1}{2k}}(t_{i+1} - t_i)}{\pi(n + 2|\Pi|k + p(n))}\right)^{2k} \leq \lambda_n(Q_{2k}) \leq \left(\frac{\sum_i (a_i^+)^{\frac{1}{2k}}(t_{i+1} - t_i)}{\pi(n - 3|\Pi|k - p(n))}\right)^{2k} \tag{19}$$

We address now the issue of the convergence of the Riemann sums. Set $I_a^\pm = \sum_i \left(a_i^\pm\right)^{\frac{1}{2k}}(t_{i+1} - t_i)$ and $I_a = \int_0^1 a^{\frac{1}{2k}} dt$. It is well known 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 [9] for and equispaced partition, we have that:

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

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 (19), adding and subtracting $\frac{I_a}{(\pi n)^{2k}}$, we find that for $n \geq \max\{n_0, k|\Pi|\}$:

$$\lambda_n(Q_{2k}) \leq \left(\frac{I_a}{\pi n}\right)^{2k} + \left(\frac{I_a^+}{\pi(n-3|\Pi|k-p(n))}\right)^{2k} - \left(\frac{I_a}{\pi n}\right)^{2k}.$$

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 $\lambda_n(Q)$. 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 (19) which decays as $n^{-2k-1/2}$ for n sufficiently large.

We have proved so far that, for one dimensional forms, Q_{2k} has $2k-$capacity $\xi_+ = (\int_0^1 \sqrt[2k]{a_t}dt)^{2k}$. Now, we apply point (i) of Proposition 3 to obtain the formula in the statement for forms on $L^2([0,1],\mathbb{R}^m)$. Finally notice that by Proposition 4 the eigenvalues of $R_k(v)$ decay as n^{-2k-1}. If we apply point (iii) of Proposition 3, we find that $Q_{2k}(v)+R_k(v)$ has the same $2k-$capacity as 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 2 and use the symmetries of 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:

$$Q(v) = \int_0^1 \langle v_k(t), \begin{pmatrix} 0 & -a_t \\ a_t & 0 \end{pmatrix} v_{k-1}(t)\rangle dt \quad a_t \geq 0 \tag{20}$$

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 $Q(Ov_0) = -Q(v_0)$ 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 us consider *the square* of Q_{2k-1}. By proposition Eq. 4 Q_{2k-1} has the same asymptotic as the form:

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

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:

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

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 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 (20); 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 $\tilde{Q}$ satisfy:

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

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 Q_{2k-1} is double and any couple $\lambda, -\lambda$ is sent to the same eigenvalue λ^2. Thus, the $(2k-1)-$capacity of Q_{2k-1} is $\left(\int_0^1 \sqrt[2k-1]{a_t}dt\right)^{2k-1}$.

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

Arguing again by point (i) of Proposition 3 one gets the estimate in the statement.

The last part about the $\infty-$capacity follow just by Proposition 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 $n\to\pm\infty$. □

4 Proof of Theorem 2

Proof of Theorem 2 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) 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) can be reformulated as $\mathcal{A}(\mathcal{V}) \subseteq \mathcal{V}^{\perp}$, thus $\mathrm{Im}(\mathcal{A}(L^2[0,1])) \subseteq \mathcal{V}^{\perp} + \mathcal{A}(\mathcal{V}^{\perp})$ which is finite dimensional.

Conversely, if the range of $\mathcal{A}$ is finite dimensional, we can decompose $L^2[0,1]$ as $\mathrm{Im}(\mathcal{A}) \oplus \ker(\mathcal{A})$, 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 (2) 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.

The first part of the statement is proved so we pass to second one. First of all notice that, now that we have written down any operator satisfying (1) and (2) in the same form as those in (3), 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 that was given in [3]. 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 $Skew(K) = \mathcal{A}$ remains 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 us forget about this change of coordinates and still call Z_t the matrix $Z_t G_t$. 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 jth component of v_i.

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

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

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 $\int_0^1 Y_t X_t^* + MX_tX_t^* dt = \int_0^1 Y_t X_t^* dt + M = 0$. The last step is to make y_j orthonormal. If we multiply Y_t by a matrix L we find the equation $L\int_0^1 Y_tY_t^* dt L^* = 1$, 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:

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

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

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

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 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 to be 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 4 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.

5 The Second Variation of an Optimal Control Problem

We start this section collecting some basic fact about optimal control problems, first and second variation. Standard references on the topic are [3, 4, 7, 10] and [8].

5.1 Symplectic Geometry and Optimal Control Problems

Consider a smooth manifold M, its cotangent bundle T^*M is a vector bundle on M whose fibre at a point q is the vector space of linear functions on T_qM, the tangent space of M at q.

Let π be the natural projection, $\pi : T^*M \to M$ which takes a covector and gives back the base point:

$$\pi : T^*M \to M, \quad \pi(\lambda_q) = q.$$

Using the the projection map we define the following $1-$form, called tautological (or Liouville) form: take an element $X \in T_\lambda(T^*M)$, $s_\lambda(X) = \lambda(\pi_* X)$. One can check that $\sigma = ds$ is not degenerate in local coordinates. We obtain a symplectic manifold considering (T^*M, σ).

Using the symplectic form we can associate to any function on T^*M a vector field. Suppose that H is a smooth function on T^*M, we define $\vec{H}$ setting:

$$\sigma(X, \vec{H}_\lambda) = d_\lambda H(X), \quad \forall X \in T_\lambda(T^*M)$$

H is called Hamiltonian function and $\vec{H}$ is an Hamiltonian vector field.

On T^*M we have a particular instance of this construction which can be used to lift arbitrary flows on the base manifold M to Hamiltonian flows on T^*M. For any vector field V on M consider the following function:

$$h_V(\lambda) = \langle \lambda, V \rangle, \quad \lambda \in T^*M.$$

It is straight forward to check in local coordinates that $\pi_* \vec{h}_V = V$.

The next objects we are going to introduce are Lagrangian subspaces. We say that a subspace W of a symplectic vector space (Σ, σ) is Lagrangian if the restriction of the symplectic form σ is degenerate, i.e. if $\{v \in \Sigma : \sigma(v, w) = 0, \forall w \in W\} = W$. An example of Lagrangian subspaces is the fibre, i.e. the kernel of π_*. More generally we can consider the following submanifolds in T^*M:

$$A(N) = \{\lambda \in T^*M : \lambda(X) = 0, \forall X \in TN, \pi(\lambda) \in N\}$$

where $N \subset M$ is a submanifold. $A(N)$ is called the annihilator of N and its tangent space at any point is a Lagrangian subspace.

Suppose we are given a family of complete and smooth vector fields f_u which depend on some parameter $u \in U \subset \mathbb{R}^k$ and a Lagrangian, i.e. a smooth function $\varphi(u, q)$ on $U \times M$. We use the vector fields f_u to produce a family of curves on M. For any function $u \in L^\infty([0, 1], U)$ we consider the following non-autonomous ODE system on M:

$$\dot{q} = f_{u(t)}(q), \quad q(0) = q_0 \in M \tag{21}$$

The solution are always Lipschitz curves. For fixed q_0, the set of functions $u \in L^\infty([0, 1], U)$ for which said curves are defined up to time 1 is an open set which we call $\mathcal{U}_{q_0}$. We can let the base point q_0 vary and consider $\mathcal{U} = \cup_{q_0 \in M} \mathcal{U}_{q_0}$. It turns out that this set has a structure of a Banach manifold (see [6]). We call the L^∞ functions obtained this way *admissible controls* and the corresponding trajectories on M *admissible curves*.

Denote by γ_u the admissible curve obtained form an admissible control u. We are interested in the following minimization problem on the space of *admissible* controls:

$$\min_{u \text{ admissible}} \mathcal{J}(u) = \min_{u \text{ admissible}} \int_0^1 \varphi(u(t), \gamma_u(t))dt \tag{22}$$

We often reduce the space of admissible variations imposing additional constraints on the final and initial position of the trajectory. For example, one can consider trajectories that start and end at two fixed points $q_0, q_1 \in M$, or trajectory that start from a submanifold N_0 and reach a second submanifold N_1. More generally we can ask that the curves satisfy $(\gamma(0), \gamma(1)) \in N \subseteq M \times M$.

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

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

We use them to lift vector fields on M to vector fields on T^*M. They are closely relate with the function defined above and still satisfy $\pi_*(\vec{h}_u) = f_u$.

In particular, if $\tilde{\gamma}$ is and admissible curve, we can build a lift, i.e. a curve $\tilde{\lambda}$ in T^*M such that $\pi(\tilde{\lambda}) = \tilde{\gamma}$, solving $\dot{\lambda} = \vec{h}_u(\lambda)$. The following theorem, known as Pontryagin Maximum Principle, gives a characterization of critical points of $\mathcal{J}$, for any set of boundary conditions.

Theorem (PMP) *If a control $\tilde{u} \in L^\infty([0, 1], U)$ is a critical point for the functional in* (22) *there exists a curve $\lambda : [0, 1] \to T^*M$ and an admissible curve $q : [0, 1] \to M$ such that for almost all $t \in [0, 1]$*

1. *$\lambda(t)$ is a lift of $q(t)$:*
$$q(t) = \pi(\lambda(t));$$
2. *$\lambda(t)$ satisfies the following Hamiltonian system:*
$$\frac{d\lambda}{dt} = \vec{h}_{\tilde{u}(t)}(\lambda);$$
3. *the control $\tilde{u}$ is determined by the maximum condition:*
$$h_{\tilde{u}(t)}(\lambda(t)) = \max_{u \in U} h_u(\lambda(t)), \quad \nu \leq 0;$$
4. *the non-triviality condition holds: $(\lambda(t), \nu) \neq (0, 0)$;*
5. *transversality condition holds:*
$$(-\lambda(0), \lambda(1)) \in A(N).$$

We call $q(t)$ an extremal curve (or trajectory) and $\lambda(t)$ an extremal.

There are essentially two possibility for the parameter ν, it can be either 0 or, after appropriate normalization of λ_t, -1. The extremals belonging to the first family are called *abnormal* whereas the ones belonging to second *normal*.

5.2 The Endpoint Map and its Differentiation

We will consider now in detail the minimization problem in equation (22) with fixed endpoints.

As in the previous section we denote by $\mathcal{U}_{q_0} \subset L^\infty([0,1], U)$ be the space of admissible controls at point q_0 and define the following map:

$$E^t : \mathcal{U}_{q_0} \to M, \quad u \mapsto \gamma_u(t)$$

It takes the control u and gives the position at time t of the solution of (21) starting from q_0. We call this map *Endpoint map*. It turns out that E^t is smooth, we are going now to compute its differential and Hessian. The proof of these facts can be found in the book [7] or in [1].

For a fixed control $\tilde{u}$ consider the function $h_{\tilde{u}}(\lambda) = h_{\tilde{u}(t)}(\lambda)$ and define the following non-autonomous flow which plays the role of parallel transport in this context:

$$\frac{d}{dt}\tilde{\Phi}_t = \vec{h}_{\tilde{u}}(\tilde{\Phi}_t) \qquad \tilde{\Phi}_0 = Id \tag{23}$$

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 $\tilde{\Phi}_t : T^*_q M \to T^*_{\tilde{\Phi}_t(q)} M$ is an affine transformation.

We suppose now that $\lambda(t)$ is an extremal and $\tilde{u}$ a critical point of the functional $\mathcal{J}$. We use the symplectomorphism $\tilde{\Phi}_t$ to pull back the whole curve $\lambda(t)$ 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:

- $b^t_u(\lambda_0)|_{u=\tilde{u}(t)} = 0 = d_{\lambda_0} b^t_u|_{u=\tilde{u}(t)}$ by definition.
- $\partial_u b^t_u|_{u=\tilde{u}(t)} = \partial_u \left(h^t_u \circ \tilde{\Phi}_t\right)|_{u=\tilde{u}(t)} = 0$ 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:

$$\begin{aligned} Z_t &= \partial_u \vec{b}^t_u(\lambda_0)|_{u=\tilde{u}(t)} : \mathbb{R}^k = T_{\tilde{u}(t)}U \to T_{\lambda_0}(T^*M) \\ H_t &= \partial^2_u b_t(\lambda_0)|_{u=\tilde{u}(t)} : \mathbb{R}^k = T_{\tilde{u}(t)}U \to T^*_{\tilde{u}(t)}U = \mathbb{R}^k \end{aligned} \tag{24}$$

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

Proposition 5 (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 its tangent space is given by:

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

When the differential of the Endpoint map is surjective a good geometric description of the situation is possible. The set of admissible control becomes smooth (at least locally) and our minimization problem can be interpreted as a constrained optimization problem. We are looking for critical points of $\mathcal{J}$ on the submanifold $\{u \in \mathcal{U} : E^t(u) = q_1\}$.

Definition 2 We say that a normal extremal $\lambda(t)$ with associated control $\tilde{u}(t)$ is strictly normal if the differential of the endpoint map at $\tilde{u}$ is surjective.

It makes sense to go on and consider higher order optimality conditions. At critical points is well defined (i.e. independent of coordinates) the Hessian of $\mathcal{J}$ (or the *second variation*). Using chronological calculus (see again [7] or [1]) it is possible to write the second variation of $\mathcal{J}$ on $\ker dE^t \subseteq L^\infty([0,1], \mathbb{R}^k)$.

Proposition 6 (Second variation) *Suppose that $(\lambda(t), \tilde{u})$ is a strictly normal critical point of $\mathcal{J}$ with fixed initial and final point. For any $u \in L^\infty([0,1], \mathbb{R}^k)$ such that $\int_0^1 X_t u_t dt = 0$ the second variation of $\mathcal{J}$ has the following expression:*

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

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

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

One often makes the assumption, which is customarily called *strong Legendre condition*, that the matrix H_t is strictly negative definite and has uniformly bounded inverse. This guarantees that the term:

$$\int_0^1 -\langle H_t u_t, v_t \rangle dt$$

is equivalent to the L^2 scalar product.

Definition 3 Suppose that the set $U \subset \mathbb{R}^k$ is open, we say that $(\lambda(t), \tilde{u})$ is a *regular* critical point if strong Legendre condition holds along the extremal. If $H_t \leq 0$ but $(\lambda(t), \tilde{u})$ does not satisfy Legendre strong condition, we say that $(\lambda(t), \tilde{u})$ is *singular*. If $H_t \equiv 0$ we say that it is *totally singular*.

Even if the extremal $(\lambda(t), \tilde{u})$ is abnormal or not strictly normal it is possible to produce a second variation for the optimal control problem. To do so, one considers the extended control system:

$$\hat{f}_{(v,u)}(q) = \begin{pmatrix} \varphi(u,q) + v \\ f_u(q) \end{pmatrix} \in \mathbb{R} \times T_q M$$

and the corresponding endpoint map $\hat{E}^t : (0, +\infty) \times \mathcal{U}_{q_0} \to \mathbb{R} \times M$. To differentiate it, we use the same construction explained above and employ the following Hamiltonians on $\mathbb{R}^* \times T^*M$:

$$\hat{h}_{(v,u)}(\nu, \lambda) = \langle \lambda, f_u \rangle + \nu(\varphi(u,q) + v)$$

One has just to identify which are the right controls to consider, PMP implies that $\dot{\nu} = 0$, $\nu \leq 0$ and $v = 0$. In the end, one obtains formally the same expression as in Proposition 6 involving the derivatives of the functions $\hat{h}_{(v,u)}$ and recover the same expression as in Proposition 6 for strictly normal extremals (see [7, Chapter 20] or [8]).

5.3 Reformulation of the Main Results

In this section, we reformulate Theorem 2 as a characterization of the compact part of the second variation of an optimal control problem at a strictly normal regular extremal (see Definitions 2 and 3).

Theorem 3 *Suppose $\mathcal{V} \subset L^2([0,1], \mathbb{R}^k)$ is a finite codimension subspace and K and operator satisfying (1) and (2). 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:*

- *a finite dimensional symplectic space (Σ, σ);*
- *a Lagrangian subspace $\Pi \subset \Sigma$;*
- *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 $\mathcal{V} = Z^{-1}(\Pi)$.

Proof The proof is essentially a reformulation of Theorem 2. 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 $\left\{\sum_{i\in I} X_{t_i}\nu_i, \nu_i \in \mathbb{R}^k, t_i \in I\right\} = \mathbb{R}^n$.

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 $\left\{\sum_{i=1}^{l} X_{t_i}\nu_i \,|\, \nu_i \in \mathbb{R}^k, l \in \mathbb{N}\right\}$ 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 $Y_t = G X_t$.

It is straightforward to check that, if $Y_t = G X_t$, 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 $\mathcal{V} = \left\{ v | \int_0^1 X_t v_t dt = 0 \right\}$. 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 (23) 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 $B_t = X_t$. □

We can reformulate also the second part of Theorem 2 relating the capacity of K and the eigenvalues of $\mathcal{A}$. We make the following assumptions:

1. the map $t \mapsto Z_t$ is piecewise analytic in t;
2. the maximum condition in the statement of PMP defines a C^2 function $\hat{H}_t(\lambda) = \max_{u \in \mathbb{R}^k} h_u^t(\lambda)$ in a neighbourhood of the strictly normal regular extremal we are considering.

Under the above assumptions the following proposition clarifies the link between the matrices Z_t and H_t and the function $\hat{H}_t$. A proof can be found either in [7, Proposition 21.3] or [1].

Proposition 7 *Suppose that $(\lambda(t), \tilde{u})$ is an extremal and the function $\hat{H}_t$ is C^2, using the flow defined in (23) define $\mathcal{H}_t(\lambda) = (\hat{H}_t - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_t(\lambda)$. It holds that:*

$$Hess_{\lambda_0}(\mathcal{H}_t) = J Z_t H_t^{-1} Z_t^* J$$

Define $R_t = \max_{v \in \mathbb{R}^k, ||v||=1} ||Z_t v||$ and let $\{\pm i \zeta_j(t)\}_{j=1}^l$ be the eigenvalues of $i Z_t^* J Z_t$ as defined in Section 4. We have the following proposition.

Proposition 8 *The capacity ξ of K satisfies:*

$$\xi \le \frac{\sqrt{k}\, ||R_t||_2}{2} \sqrt{\int_0^1 tr(Hess_{\lambda_0}(\mathcal{H}_t)) dt}$$

and in particular, if we arrange the functions $\zeta_j(t)$ in a decreasing order, they satisfy

$$0 \le \zeta_j(t) \le R_t \sqrt{\lambda_{2j}(t)}, \quad j \in \{1, \dots l\}$$

where $\lambda_j(t)$ are the eigenvalues of $Hess_{\lambda_0}(\mathcal{H}_t)$ in decreasing order.

Proof We give a sketch of the proof. Without loss of generality we can assume $H_t = -Id$, otherwise, we can perform the change of coordinate on $L^2([0,1],\mathbb{R}^k)$ $v \mapsto (-H_t)^{-\frac{1}{2}}v$ and redefine Z_t accordingly.

In this notation $Hess_{\lambda_0}(\mathcal{H}_t)$ corresponds to the matrix $JZ_tZ_t^*J$. If we square $A_t = Z_t^*JZ_t$ we obtain:

$$A_t^*A_t = -Z_t^*JZ_tZ_t^*JZ_t = -Z_t^*\left(JZ_tZ_t^*J\right)Z_t = -Z_t^*Hess_{\lambda_0}(\mathcal{H}_t)Z_t$$

Observe that $\zeta_j(t)$ is an eigenvalue of A_t if and only if $-\zeta_j^2(t)$ is a eigenvalue of $A_t^*A_t$. The equation above relates the *restriction* of $Hess_{\lambda_0}(\mathcal{H}_t)$ to the image of the maps $Z_t : \mathbb{R}^k \to \mathbb{R}^{2n}$ with the square of the functions $\zeta_j(t)$ defining the capacity.

The idea is to use Cauchy interlacing inequality for the eigenvalues of $Hess_{\lambda_0}(\mathcal{H}_t)$ and its restriction to a codimension $2n-k$ subspace. If $\{\lambda_j(t)\}_{j=1}^{2n}$ are the eigenvalues of the Hessian, taken in decreasing order, and $\{\mu_j(t)\}_{j=1}^{2n-k}$ the eigenvalues of its restriction we have:

$$\lambda_{j+2n-k}(t) \le \mu_j(t) \le \lambda_j(t)$$

In our case, Z_t are not orthogonal projectors but we can adjust the estimates considering how much the matrices Z_t dilate the space, and thus we have to take in account the function R_t defined just before the statement. Denote by $\mu_j(t)$ the jth eigenvalue of $-A_t^2$, putting all together we have:

$$0 \le \mu_j(t) \le R_t^2\lambda_{2j}(t) \quad j \in \{1,\dots k\}$$

where we shifted the index by one since $\mu_{2k-1}(t) = \mu_{2k}(t)$ for all $k \le l$. Taking square roots and integrating we have:

$$\int_0^1 \zeta_j(t)dt \le \int_0^1 R_t\sqrt{\lambda_{2j}(t)}dt$$

Summing up over j we find that:

$$\xi = \int_0^1 \sum_j \zeta_j(t)dt \le \frac{1}{2}\int_0^1 \sum_j R_t\sqrt{\lambda_{2j}(t)}dt \le \frac{\sqrt{k}||R_t||_2}{2}\sqrt{\int_0^1 \mathrm{tr}(Hess_{\lambda_0}(\mathcal{H}_t))}$$

□

We turn now to Theorem 1; we can interpret it as a quantitative version of various necessary optimality conditions that one can formulate for certain classes of singular extremals (see [7, Chapter 20] or [4, Chapter 12]). Moreover, leaving optimality conditions aside, Theorem 1 gives the asymptotic distribution of the eigenvalues of the second variation for totally singular extremals (see definition 3).

As mentioned in the previous section, we can produce a 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:

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

In this case, Legendre condition $H_t \leq 0$ (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 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 1, Goh conditions are written as $Q_1 \equiv 0$ 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, it is clear that if $Q_1 \not\equiv 0$, 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 Q_2 is non-positive definite (recall the different sign convention in Proposition 6); otherwise, the negative part of the spectrum of $-Q_2$ 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 $Q_2 \equiv 0$, one can take more derivatives and find new conditions. In particular, using the notation of Lemma 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.

Acknowledgements The author wishes to thank Prof. A. Agrachev for the stimulating discussions on the topic and the referee for the helpful suggestions which greatly improved the exposition.

Funding Open access funding provided by Scuola Internazionale Superiore di Studi Avanzati - SISSA within the CRUI-CARE Agreement.

References

1. Agrachev A, Stefani G, Zezza P. An invariant second variation in optimal control. Internat J Control. 1998;71(5):689–715.
2. Agrachëv AA. Quadratic mappings in geometric control theory. In: Problems in geometry, Vol. 20 (Russian). Itogi Nauki i Tekhniki. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1998. Translated in J. Soviet Math.; 1988. p. 111–205. **5**1 (1990), no. 6, 2667–2734.
3. Agrachev AA. Spectrum of the second variation. Tr Mat Inst Steklova. 2019;304(Optimal noe Upravlenie i Differentsial nye Uravneniya):32–48.
4. Agrachev A, Barilari D, Boscain U, vol. 181. A comprehensive introduction to sub-Riemannian geometry, Cambridge studies in advanced mathematics. Cambridge: Cambridge University Press; 2020. From the Hamiltonian viewpoint, With an appendix by Igor Zelenko.
5. Agrachev A, Beschastnyi I. Jacobi fields in optimal control: one-dimensional variations. J Dyn Control Syst. 2020;26(4):685–732.

6. Agrachev A, Beschastnyi I. Jacobi fields in optimal control: Morse and Maslov indices. Nonlinear Anal, 214. Paper No. 112608, 47. 2022.
7. Agrachev AA, Sachkov YL. Control theory from the geometric viewpoint. In: Encyclopaedia of mathematical sciences. Berlin: Springer; 2004. Control Theory and Optimization, II.
8. Agrachev AA, Beschastnyi IY. Symplectic geometry of constrained optimization. Regul Chaotic Dyn. 2017;22(6):750–770.
9. Chui CK. Concerning rates of convergence of Riemann sums. J Approximation Theory. 1971;4:279–287.
10. Jean F. Control of nonholonomic systems: from sub-Riemannian geometry to motion planning, SpringerBriefs in Mathematics. Cham: Springer; 2014.
11. Kato T. Perturbation theory for linear operators, Classics in mathematics. Berlin: Springer; 1995. Reprint of the 1980 edition.
12. Rudin W. Functional analysis, 2nd ed., International series in pure and applied mathematics. New York: McGraw-Hill, Inc.; 1991.