Port. Math. 81 (2024), 107–126
DOI 10.4171/PM/2117

Morse index of block Jacobi matrices via optimal control

Stefano Baranzini and Ivan Beschastnyi

Abstract. We describe the relation between block Jacobi matrices and minimization problems for discrete-time optimal control problems. Using techniques developed for the continuous case, we provide new algorithms to compute and estimate spectral invariants of block Jacobi matrices.

1. Introduction

The goal of this paper is to explore an interesting connection between block Jacobi matrices and a class of discrete optimal control problems. This gives effective algorithms for computing the negative inertia index and more generally the number of eigenvalues smaller than some threshold $\lambda^* \in \mathbb{R}$ of large block Jacobi matrices.

Recall that a block Jacobi matrix $\mathcal{J}$ is a matrix of the form

$$\mathcal{J} = \begin{pmatrix} S_1 & R_1 & 0 & \dots & 0 & 0 \\ R_1^* & S_2 & R_2 & \dots & 0 & 0 \\ 0 & R_2^* & S_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & S_{N-1} & R_{N-1} \\ 0 & 0 & 0 & \dots & R_{N-1}^* & S_N \end{pmatrix}, \tag{1.1}$$

where $N \in \mathbb{N}$, S_i are Hermitian matrices of order n, R_i is a complex matrix and R_i^* denotes the conjugate transpose of R_i.

Jacobi matrices find applications in numerical analysis [6], statistical physics [7], knot theory [5] and many other areas. Moreover, any Hermitian matrix can be put in a tridiagonal form using Householder transformations [4]. Therefore, understanding their spectral properties is an extremely important topic.

In this article, we establish a correspondence between Jacobi matrices and the discrete linear quadratic regulator (LQR) problem. We combine optimal control theory with matrix theory to establish ways of computing or estimating the Morse index

Mathematics Subject Classification 2020: 15B99 (primary); 49N10, 15A18 (secondary).
Keywords: Jacobi matrix, discrete systems, Morse index.

of such matrices and of the second variation of discrete optimal control problems. The techniques used in our proofs are deeply connected with symplectic geometry. One can extend them to much more general settings [1–3]. More precisely, exploiting this connection, we prove a formula linking the negative inertia index $\operatorname{ind}^-(\mathcal{I}-\lambda)$ (i.e., the number of *negative* eigenvalues of $\mathcal{I}-\lambda$) and the fundamental solution of a suitable discrete time differential equation.

As a first application, we prove a slight generalization of [12, Theorem 10.1]. See also [10, 11] for a detailed analysis from an intersection theory perspective. Assume that $R_i \in \mathrm{GL}(n,\mathbb{C})$ for all $i \in \{1,\dots,N-1\}$. Let $S_0 = S_0^*$ be any Hermitian matrix and $R_0, R_N \in \mathrm{GL}(n,\mathbb{C})$. Define the following matrices:

$$\begin{aligned} M_k(\lambda) &:= \begin{pmatrix} -R_k^{-1} & R_k^{-1}(1-S_k+\lambda+R_kR_k^*) \\ -R_k^{-1} & R_k^{-1}(1-S_k+\lambda) \end{pmatrix}, \\ \Phi_k(\lambda) &:= \prod_{j=0}^{k-1} M_j(\lambda) = \begin{pmatrix} \Phi_k^1(\lambda) & \Phi_k^2(\lambda) \\ \Phi_k^3(\lambda) & \Phi_k^4(\lambda) \end{pmatrix}. \end{aligned} \tag{1.2}$$

We shall prove the following theorem.

Theorem 1. *Let $\mathcal{I}$ be a Jacobi matrix as in (1.1) and let $R_i \in \mathrm{GL}(n,\mathbb{C})$ for all $i \in \{1,\dots,N-1\}$. The negative inertia index of $\mathcal{I}-\lambda$ is given by the following formula:*

$$\operatorname{ind}^-(\mathcal{I}-\lambda) = \sum_{k=1}^{N} \operatorname{ind}^- \mathcal{H}_k^\lambda + \dim\ker\Phi_k^3(\lambda), \tag{1.3}$$

where

$$\mathcal{H}_k^\lambda = -(\Phi_k^3(\lambda))^* R_k \Phi_{k+1}^3(\lambda).$$

Remark 1. It should be noted that $\mathcal{H}_k^\lambda$ are indeed Hermitian and hence the negative inertia index is well defined.

Remark 2. Notice that the matrices $M_k(\lambda)$ appearing in the statement of Theorem 1 can be rewritten in a more standard form as transfer matrices. Indeed, set

$$T_k(\lambda) = \begin{pmatrix} -(S_k-\lambda)R_{k-1}^{-1} & -R_{k-1}^* \\ R_{k-1}^{-1} & 0 \end{pmatrix}, \qquad G_k = \begin{pmatrix} 0 & R_k \\ (R_k^{-1})^* & -(R_k^{-1})^* \end{pmatrix}.$$

A straightforward computation yields $M_k(\lambda) = G_k^{-1}T_k(\lambda)G_{k-1}$. Moreover, the matrix $\Psi_{k+1}(\lambda) := G_k\Phi_{k+1}(\lambda)$ coincides with the fundamental solution used in [12]. A straightforward computation shows that

$$\mathcal{H}_k^\lambda = -(\Psi_{k+1}^3(\lambda))^* R_{k+1}\Psi_{k+2}(\lambda) \qquad \text{and} \qquad \ker\Phi_k^3(\lambda) = \ker\Psi_{k+1}^3(\lambda).$$

Let us consider now the case in which the R_i are not necessarily invertible but let us assume that $R_i = R_i^*$. This is not restrictive since a similarity transformation reduces $\mathcal{I}$ to this case (see, for instance, [12, Section 9]). Since the R_i are Hermitian, $\mathbb{C}^n$ splits as a direct sum of $\ker R_i$ and its orthogonal, $\operatorname{Im} R_i$. Denote by π_i the projection onto $(\ker R_i)^\perp$. Let L_i be an $n \times n$ matrix such that $R_i + L_i$ is invertible and Hermitian and $R_i L_i = 0$. Finally, let again $S_0 = S_0^*$ be any Hermitian matrix and $R_0, R_N \in \mathrm{GL}(n, \mathbb{C})$. Let us define the following matrices:

$$M_k(\lambda) = \begin{pmatrix} -L_k^\dagger - R_k^\dagger & (L_k^\dagger + R_k^\dagger)(\pi_{k-1} - S_k + \lambda) + R_k \\ -R_k^\dagger & R_k^\dagger(\pi_{k-1} - S_k + \lambda) - L_k \end{pmatrix}, \tag{1.4}$$

and let $\Phi_k(\lambda)$ be defined as in equation (1.2). Here $L_k^\dagger$ and $R_k^\dagger$ are the Moore–Penrose pseudo-inverses of the corresponding matrices.

Theorem 2. *Let $\mathcal{I}$ be a Jacobi matrix as in (1.1) and let R_i be Hermitian (and possibly degenerate) matrices for all $i \in \{1, \dots, N-1\}$. The following bound holds:*

$$\operatorname{ind}^-(\mathcal{I} - \lambda) \geq \sum_{k=1}^{N} \operatorname{ind}^- \mathcal{H}_k^\lambda + \dim\big(\ker \Phi_k^3(\lambda) / (\ker \Phi_{k+1}^3(\lambda) \cap \ker \Phi_k^3(\lambda))\big), \tag{1.5}$$

where the quadratic form $\mathcal{H}_k^\lambda$ is defined as follows:

$$\mathcal{H}_k^\lambda(\eta_0) = -\langle (\Phi_k^3(\lambda))^* R_k (\Phi_{k+1}^3(\lambda) + (1 - \pi_{k-1})\Phi_k^1(\lambda))\eta_0, \eta_0 \rangle,$$

where $\eta_0 \in \ker(1 - \pi_k)\Phi_k^3(\lambda)$.

Firstly, it should be noted that Theorem 2 reduces to Theorem 1 when all R_k are invertible and the inequality actually becomes equality. Secondly, the lower bound in (1.5) is not necessarily the best one obtainable applying Theorem 3 and depends on the particular choice of $M_k(\lambda)$ adopted in (1.4). The reason is that the right-hand side of (1.5) corresponds to the index of $\mathcal{I}$ restricted to a specific subspace determined by the choice of (1.4). There is some freedom in this choice: to solutions of (3.3) (up to some equivalence) correspond distinct subspaces, as explained in Section 3 and Remark 5.

The paper has the following structure. In Section 2, we first prove Theorem 1 using the technique of transfer matrices. This will give an idea for the proof of the more general Theorem 3 from which both, Theorems 1 and 2, follow directly. In Section 3, we discuss the discrete LQR problem and its connection to Jacobi matrices. In Section 4, we compute the index of quadratic forms associated to LQR problems. In Section 4.1, we derive the two theorems above from our formula.

2. Proof of Theorem 1 via transfer matrices

In order to prove Theorem 1 and Theorem 2, we will exploit in Sections 3 and 4 the connection between Jacobi matrices and optimal control. This will allow us to obtain proofs of both theorems as a consequence of a more general index formula for LQR problems in a single unified manner. However, Theorem 1 was already proven in [12] using the technique of transfer matrices and a discrete version of the Sturm–Liouville theory. The author of [12] uses mathematical physics as a guiding argument, while we use optimal control theory. In the intersection of the two approaches lies a purely algebraic argument that is instructive for understanding both proofs.

Assume that the blocks R_j of the Jacobi matrix $\mathcal{I}$ in (1.1) are invertible. To compute the Morse index we rely on the following lemma.

Lemma 1 ([8, Corollary 2.7]). *Suppose that a quadratic form $\mathcal{Q}$ is defined on a finite dimensional vector space X and a subspace $W \subseteq X$ is given. Denote by $W^{\perp_{\mathcal{Q}}}$ the $\mathcal{Q}$-orthogonal subspace to W, i.e.,*

$$W^{\perp_{\mathcal{Q}}} := \{u \in X : \mathcal{Q}(u,v) = 0, \forall v \in W\}.$$

Denote by $\operatorname{ind}^-\mathcal{Q}$ *the number of negative eigenvalues of $\mathcal{Q}$, then*

$$\operatorname{ind}^-\mathcal{Q} = \operatorname{ind}^-\mathcal{Q}|_W + \operatorname{ind}^-\mathcal{Q}|_{W^{\perp_{\mathcal{Q}}}} + \dim\big((W \cap W^{\perp_{\mathcal{Q}}})/(W \cap W^{\perp_{\mathcal{Q}}} \cap \ker \mathcal{Q})\big).$$

Remark 3. Here is an idea of the proof. If both $\mathcal{Q}$ and the restriction $\mathcal{Q}|_W$ are non-degenerate, then there is no dimension term and the result is standard. One can prove this using, for example, a Gram–Schmidt argument. Factoring out $\ker \mathcal{Q} \subseteq W \cap W^{\perp_{\mathcal{Q}}}$, one can assume that $\mathcal{Q}$ is not degenerate while $\mathcal{Q}|_W$ is. One then notices that $W \cap W^{\perp_{\mathcal{Q}}}$ is isotropic for both $\mathcal{Q}|_W$ and $\mathcal{Q}|_{W^{\perp_{\mathcal{Q}}}}$. Moreover, $\mathcal{Q}$ induces a non-degenerate pairing on

$$X/(W + W^{\perp_{\mathcal{Q}}}) \oplus (W \cap W^{\perp_{\mathcal{Q}}})$$

which has dimension $2\dim(W \cap W^{\perp_{\mathcal{Q}}})$ yielding the result. More details can be found, for instance, in [8] or in [2, Appendix B].

Let us define the following family of subspaces:

$$\mathcal{W}_k = \{x \subseteq \mathbb{R}^{nN} : x_i = 0, \forall i \geq k+1\},$$

which are just vectors in $\mathbb{R}^{nN}$ having zero component starting from step $Nk+1$. We will iteratively apply Lemma 1 choosing $X = \mathcal{W}_{k+1}$, $\mathcal{Q}(x) = \langle x, \mathcal{I}x\rangle$ for $x \in X$ and $\mathcal{W}_k$ as W. Clearly $\mathcal{W}_{k-1} \subset \mathcal{W}_k$ and $\bigcup_{k=1}^{N+1} \mathcal{W}_k = \mathbb{R}^{nN}$. Thus the $\mathcal{W}_k$ provide a filtration of $\mathbb{R}^{nN}$.

Let us introduce some simplified notation. Let $\mathcal{I}_k$ be be the quadratic form induced on $\mathcal{W}_k$ by the principal $nk \times nk$ minor of $\mathcal{I}$. Set

$$\mathcal{W}_k^{\perp} := \mathcal{W}_k^{\perp_{\mathcal{I}_{k+1}}}, \qquad \mathcal{I}_k^{\perp} := \mathcal{I}_k|_{\mathcal{W}_{k-1}^{\perp}}.$$

Note that by definition $\mathcal{W}_k^{\perp} \subset \mathcal{W}_{k+1}$. Moreover, we have that

$$\mathcal{I}_{k+1}|_{\mathcal{W}_k} = \mathcal{I}_k.$$

Let us apply Lemma 1 with $\mathcal{I} = \mathcal{I}_{k+1}$ and $W = \mathcal{W}_k \subseteq \mathcal{W}_{k+1} = X$. We obtain

$$\operatorname{ind}^- \mathcal{I}_{k+1} = \operatorname{ind}^- \mathcal{I}_k + \operatorname{ind}^- \mathcal{I}_{k+1}^{\perp} + \dim\big((\mathcal{W}_k \cap \mathcal{W}_k^{\perp})/(\mathcal{W}_k \cap \mathcal{W}_k^{\perp} \cap \ker \mathcal{I}_{k+1})\big).$$

Thus, by iteration, we obtain

$$\begin{aligned}\operatorname{ind}^- \mathcal{I} = \operatorname{ind}^- \mathcal{I}_1 + \sum_{k=2}^{N} & \operatorname{ind}^- \mathcal{I}_k^{\perp} \\ & + \dim\big((\mathcal{W}_{k-1} \cap \mathcal{W}_{k-1}^{\perp})/(\mathcal{W}_{k-1} \cap \mathcal{W}_{k-1}^{\perp} \cap \ker \mathcal{I}_k)\big).\end{aligned}$$

Now we compute each $\mathcal{I}_k^{\perp}$. First we describe elements x which belong to $\mathcal{W}_k^{\perp}$. Testing against elements of $\mathcal{W}_k$, we find that $\langle x, \mathcal{I} y\rangle = 0$ when $x \in \mathcal{W}_{k+1}$ for all $y \in \mathcal{W}_k$ if and only if

$$\begin{cases} S_1 x_1 + R_1 x_2 = 0, & \\ S_j x_j + R_{j-1}^* x_{j-1} + R_j x_{j+1} = 0, & \text{for } 1 < j \le k, \\ x_j = 0, & \text{for } j > k+1. \end{cases}$$

We can extend the first equation to include also the vector x_0 requiring

$$R_0 x_0 + S_1 x_1 + R_1 x_2 = 0.$$

We require that $x_0 = 0$ and set, without loss of generality, $R_0 := 1$. In this way, we get the same formula for all indices, including the first one. From those equations we obtain for $0 \le j \le k+1$

$$\begin{pmatrix} R_j x_{j+1} \\ x_j \end{pmatrix} = \begin{pmatrix} -S_j x_j - R_{j-1}^* x_{j-1} \\ x_j \end{pmatrix} =: T_j \begin{pmatrix} R_{j-1} x_j \\ x_{j-1} \end{pmatrix}$$

where

$$T_j = \begin{pmatrix} -S_j R_{j-1}^{-1} & -R_{j-1}^* \\ R_{j-1}^{-1} & 0 \end{pmatrix}$$

are the *transfer matrices*. We can define a flow as

$$\Psi_k := \prod_{j=1}^{k} T_j. \tag{2.1}$$

Let Ψ_k^3 be the lower left $n \times n$ block. Now we are left with computing $\mathcal{I}_{k+1}(x)$ for $x \in \mathcal{W}_k^\perp$,

$$\mathcal{I}_k^\perp(x) = \langle x, \mathcal{I}x\rangle = \langle x_{k+1}, S_{k+1}x_{k+1} + R_k^* x_k\rangle = -\langle \Psi_{k+1}^3 x_1, R_{k+1}\Psi_{k+2}^3 x_1\rangle.$$

Hence, the negative index of $\mathcal{I}_k$ coincides with the negative Morse index of the matrix $-(\Psi_{k+1}^3)^* R_{k+1}\Psi_{k+2}^3$.

It only remains to compute the dimensions of intersections. It is clear from the definitions of $\mathcal{W}_k$ that

$$\mathcal{W}_k^\perp \cap \mathcal{W}_k = \ker \mathcal{I}_k;$$
$$\mathcal{W}_k^\perp \cap \mathcal{W}_k \cap \ker \mathcal{I}_{k+1} = \ker \mathcal{I}_k \cap \ker \mathcal{I}_{k+1}.$$

It is straightforward to check that elements in $\mathcal{W}_k^\perp \cap \mathcal{W}_k$ satisfy the additional requirement that $x_{k+1} = 0$. Thus $\ker \mathcal{I}_k$ is isomorphic to $\ker \Psi_{k+1}^3$. Moreover, a direct computations shows that

$$\ker \Psi_{k+1}^3 \cap \ker \Psi_{k+2}^3 = (0).$$

Collecting all the pieces gives a formula analogous to (1.3) in Theorem 1. In order to recover exactly the statement of Theorem 1 it is enough to apply Remark 2.

Remark 4. What is the difference between the flow Ψ defined in (2.1) and Φ as given in (1.2)? The map $x \mapsto \langle x, Ix\rangle$ can be viewed as a discrete version of an action functional. In the continuous case, the corresponding Euler equations for critical points are a system of second order ODEs. One can rewrite it as a system of first order equations double in size. But this passage is not unique. The flows Φ and Ψ can be thought as fundamental solutions of the discretization of the extremal equations for two different choices. Because of this, it is not surprising that in the end, we compute the same object in two slightly different ways. There is, however, a particular choice using the Legendre transform. It gives naturally the extremal equations a structure of a Hamiltonian system. Our flow Φ corresponds exactly to this case. It should be noted that having a Hamiltonian description, as given by Pontryagin's maximum principle, is particularly useful when dealing with constrained variational problems. This allows us to treat also the more singular situation stated in Theorem 2.

3. LQR and Jacobi matrices

3.1. Linear quadratic problems

In this section, we introduce all the necessary notations and definitions concerning optimal control problems (OCPs for short) and then, we establish the aforementioned correspondence with Jacobi matrices. More precisely, we formulate second order optimality conditions in terms of a suitable Jacobi matrix determined by the OCP under consideration.

In the following, we will consider LQR problems, see [9, Chapter 2] for further references. They are discrete time OCP having a linear *dynamic* and a quadratic *cost* functional. The reasons for such a choice are twofold. Firstly, LQR system are far simpler to handle than general *non-linear* optimal control problems. Secondly, once an extremal of a possibly non-linear optimal control problem is fixed, second order optimality conditions can be formulated in terms of a (non-autonomous) LQR problem, i.e., the *linearization* of the system along the extremal.

Let us introduce the *dynamic* and the space of trajectories on which we are going to minimize. For $0 \leq k \leq N+1$, let A_k be $n \times n$ matrices, and B_k be $n \times m_k$ matrices with $m_k \leq n$. Moreover, let $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^{m_k}$. Consider the following difference equation:

$$x_{k+1} - x_k = A_k x_k + B_k u_k, \qquad k = 0, \dots, N+1, \tag{3.1}$$

For technical reasons, throughout the rest of the paper, we will make the following assumption.

Assumption 1. *The matrices* $1 + A_k$ *in* (3.1) *are invertible for every* k. *The matrices* B_k *are injective for* $1 \leq k \leq N+1$ *and* B_0 *is invertible.*

The variables x_k are called *state* variables while u_k *controls*. We denote by $x = (x_k)$ and $u = (u_k)$ the vectors having x_k and u_k as k-th component. Note that we will often consider them as elements of $\mathbb{R}^{n(N+2)}$ and $\mathbb{R}^{\sum_{k=0}^{N+1} m_k}$ respectively. To any fixed initial condition x_0 and any control u corresponds a unique solution $x(u, x_0)$ of equation (3.1). We will refer to it as an *admissible trajectory*.

Let us consider the following quadratic *cost* functional on the space of controls:

$$\widetilde{\mathcal{J}}(u) = \frac{1}{2} \sum_{i=0}^{N} \left(|u_i|^2 - \langle Q_i x_i(u, x_0), x_i(u, x_0) \rangle \right), \tag{3.2}$$

where Q_i are symmetric $n \times n$ matrices. Note that one can equivalently work on the space of solutions $x(u, x_0)$ with fixed initial condition x_0. It is straightforward to check that these spaces are isomorphic since the B_k are injective. Thus, it is possible to rewrite (3.2) as a function on the space of solutions of (3.1). The problem of minimizing $\widetilde{\mathcal{J}}$ over the latter space is the classical linear quadratic regulator problem, which is one of the central problems in optimal control theory.

Let us describe here explicitly the correspondence between u and solutions x of (3.1), i.e., between *controls* and *admissible trajectories*. Recall that by the properties of the pseudo-inverse, $B_k B_k^\dagger$ is the orthogonal projection (with respect to the standard scalar product) on $\operatorname{Im}(B_k) \subseteq \mathbb{R}^n$. It follows that x is a solution of equation (3.1) if and only if

$$(1 - B_k B_k^{\dagger})(x_{k+1} - (1 + A_k)x_k) = 0, \qquad \forall 0 \leq k \leq N + 1,$$

which equivalently says that $x_{k+1} - (1 + A_k)x_k$ is zero modulo $\operatorname{Im} B_k$.

Let us denote the space of admissible trajectories as $\mathcal{W}$. Inverting relation (3.1) and combining it with (3.2) we obtain a quadratic form $\mathcal{J}$ on $\mathcal{W}$,

$$\mathcal{J}(x) = \langle \mathcal{I} x, x \rangle = \frac{1}{2}\Bigg(\sum_{i=0}^{N} |B_i^{\dagger}(x_{i+1} - (1 + A_i)x_i)|^2 - \sum_{i=0}^{N} \langle Q_i x_i, x_i \rangle \Bigg), \qquad x \in \mathcal{W}.$$

Throughout the rest of the paper we will impose *homogeneous Dirichlet* boundary conditions, i.e., we will additionally assume that $x_0 = 0 = x_{N+1}$. Let us write down the matrix $\mathcal{I}$ associated to the quadratic form above; it is given by a block Jacobi matrix of the form (1.1). Define $\Gamma_i = B_i B_i^t \geq 0$ which is a symmetric and non-negative $n \times n$ matrix. After some algebraic manipulation, we find the following relations:

$$\begin{aligned} S_i &= \Gamma_{i-1}^{\dagger} + (1 + A_i)^t \Gamma_i^{\dagger} (1 + A_i) - Q_i, \\ R_i &= -(1 + A_i)^t \Gamma_i^{\dagger}, \end{aligned} \tag{3.3}$$

$1 \leq i \leq N - 1$.

It follows that to each LQR problem (3.1)–(3.2) we can associate a couple $(\mathcal{I}, \mathcal{W})$ where $\mathcal{I}$ is a real Jacobi matrix and $\mathcal{W}$ a subspace of $\mathbb{R}^{n(N+1)}$. The converse is also true. However, there are several LQR problems which correspond to the same Jacobi matrix.

Remark 5. Contrary to the scalar case, when the R_i are degenerate (but non-zero), general formulas to compute index and determinant of block Jacobi matrices are not known. In this case, in fact, it is not possible to define the usual *transfer matrices* and $\mathcal{I}$ does not decompose as the sum of smaller matrices sharing the same block structure. Our approach has the following consequence: since for the subspaces $\mathcal{W}$ of *admissible trajectories* we are able to define a symplectic flow, similar to the one defined when $\det R_i \neq 0$, and since we always have that

$$\operatorname{ind}^- \mathcal{I} \geq \operatorname{ind}^- \mathcal{I}|_{\mathcal{W}}, \tag{3.4}$$

we can prove inequalities of the form (1.5) expressing the right-hand side of (3.4) in terms of the symplectic flow we build. Note that on the one hand, the estimates might worsen as the kernel of the R_i enlarges but on the other hand we have a lot of freedom in choosing $\mathcal{W}$.

3.2. Jacobi matrices and Lagrange multiplier rule

In order to compute the index of a Jacobi matrix $\mathcal{I}$ in (1.1), we will use the flow of a certain system of difference equations. First, though, we need some notions from optimal control theory.

Let $U := \mathbb{R}^{mN}$, the *space of controls*. The first object we introduce is the *endpoint map*

$$E^{k+1} : U \to \mathbb{R}^n, \quad u \mapsto x_{k+1}(u), \qquad 0 \le k \le N.$$

This map takes a control u and gives the solution of the differential equation (3.1) with $x_0 = 0$ at step $k+1$. For brevity define the following matrices for $0 \le i \le j \le N+1$:

$$P_i^j := \prod_{r=i}^{j-1}(1 + A_r), \quad P_j^j = 1,$$

and for $0 \le i < j \le N+1$ we take

$$P_j^i = (P_i^j)^{-1}.$$

We can write the endpoint map explicitly iterating (3.1), namely

$$E^{k+1}(u) = x_{k+1}(u_0, \dots, u_N) = \sum_{j=1}^{k+1} P_j^{k+1} B_{j-1} u_{j-1}. \tag{3.5}$$

In particular, from this formula, it is clear that the value of $E^{k+1}(u)$ depends only on the first k components of the control u. It makes sense, thus, to introduce the following filtration (i.e., *flag* of subspaces) in the space of controls U. For $0 \le k \le N$ let us define

$$U_k := \{(u_0, u_1, \dots, u_k, 0, \dots, 0) : u_j \in \mathbb{R}^m\}, \quad U_N = U. \tag{3.6}$$

We will often identify U_k with a copy of $\mathbb{R}^{m(k+1)}$ and suppress the extra zeros in equation (3.6) to simplify notation. Denote by pr_l the orthogonal projection on U_l with respect to the standard Euclidean scalar product. Namely a linear operator $\mathrm{pr}_l : U \to U$ defined as:

$$\mathrm{pr}_l(u) = \begin{cases} u, & u \in U_l, \\ 0, & u \in U_l^{\perp}. \end{cases}$$

It is straightforward to check that

$$E^{k+1}(u) = E^{k+1}(\mathrm{pr}_l u) \qquad \forall l \ge k.$$

Lemma 2. *Under Assumption 1, the differential of the endpoint map is surjective. For $1 \leq k \leq N-1$, consider the maps*

$$F_k : U_{k-1} \to \mathbb{R}^m,$$

defined as

$$F_k(u) := -B_k^{\dagger} \sum_{j=1}^{k} P_j^{k+1} B_{j-1} u_{j-1}.$$

Then $(u_0, \dots, u_k) \in \ker E^{k+1} \cap U_k$ if and only if

$$u_k = F_k(u) \qquad \text{and} \qquad (1 - B_k B_k^{\dagger}) P_k^{k+1} E^k(u) = 0. \tag{3.7}$$

Proof. To show that the differential of the endpoint map is surjective, it is enough to find a subspace that maps isomorphically to $\mathbb{R}^n$. For controls of the form $u = (u_0, 0, 0, \dots, 0)$, under Assumption 1, we have $E^{k+1}(u) = P_1^{k+1} B_0 u_0$ which clearly is onto.

The proof of the second part of the statement follows by a simple algebraic manipulation of equation (3.5). Whenever B_k is not invertible the extra condition in equation (3.7) appears. Notice that $1 - B_k B_k^{\dagger}$ is the orthogonal projection on $\operatorname{Im}(B_k)^{\perp}$, this component cannot be modified using just u_k. ∎

The following result is a discrete version of the Pontryagin maximum principle (PMP) for LQR problems, which characterizes critical points of the constrained variational problem as solutions of a Hamiltonian system.

Lemma 3. *The quadratic form $\widetilde{\mathcal{J}}$ is degenerate on $\ker E^{N+1}$ if and only if the boundary value problem*

$$\begin{cases} x_{k+1} = P_k^{k+1} x_k + B_k B_k^t \lambda_{k+1}, \\ \quad \lambda_k = (P_k^{k+1})^t \lambda_{k+1} + Q_k x_k, \end{cases} \tag{3.8}$$

with $x_0 = 0$, $x_{N+1} = 0$ has a non-trivial solution. Moreover, the dimension of the space of solutions of equation (3.8) and of $\ker \widetilde{\mathcal{J}}|_{\ker E^{N+1}}$ coincide.

Proof. We will employ the Lagrange multipliers rule to write down the equation for critical points of $\widetilde{\mathcal{J}}$ restricted to $\ker E^{N+1}$. Recall that, given two smooth functions $f : \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}^s$, finding critical points of f on the level set $g^{-1}(c)$ (when c is a regular value) is equivalent to finding a solution $(x, \lambda) \in \mathbb{R}^{d+s}$ of the equations

$$\begin{cases} g(x) = c, \\ d_x f = \lambda^t \, d_x g. \end{cases}$$

We will construct a discrete Hamiltonian system using the filtration (3.6) and applying iteratively the Lagrange multipliers rule. To this extent, we need to introduce a suitable family of functionals $\widetilde{\mathcal{J}}^k$, defined on U_k. They play the role of the *restriction* of $\widetilde{\mathcal{J}}$ to U_k, even if they differ from it slightly. For $1 \le k \le N$ and $u \in U_k$ let us define

$$\widetilde{\mathcal{J}}^k(u) := \frac{1}{2}\Bigg(\sum_{i=0}^{k} |u_i|^2 - \sum_{i=1}^{k}\langle x_i, Q_i x_i\rangle\Bigg). \tag{3.9}$$

Their differentials at a point u are then given by

$$d_u\widetilde{\mathcal{J}}^k(v) = \sum_{i=0}^{k}\langle u_i, v_i\rangle - \sum_{i=1}^{k}(E^i(u))^t Q_i E^i(v). \tag{3.10}$$

Notice that $\widetilde{\mathcal{J}}^N = \widetilde{\mathcal{J}}$. As already discussed, if u is a critical point of $\widetilde{\mathcal{J}}$ restricted to $\ker E^{N+1}$, there exists $\lambda_{N+1} \in \mathbb{R}^n$ such that

$$d_u\widetilde{\mathcal{J}}^N = \lambda_{N+1}^t E^{N+1}. \tag{3.11}$$

We now look for $\lambda_k \in \mathbb{R}^n$, $1 \le k \le N+1$, which satisfy

$$d_u\widetilde{\mathcal{J}}^k = \lambda_{k+1}^t E^{k+1}. \tag{3.12}$$

We already have an equation for the trajectory $(0, x_1, \ldots, x_N, 0)$ determined by a critical point u, this is given by (3.1). We still need to find a relation linking the controls u_k with the covectors λ_k^t and a difference equation for the covectors λ_k^t. Let us address the first task. Using the explicit formula (3.5) for the endpoint map, we find the recurrence relation

$$E^{k+1}(v) = P_k^{k+1}E^k(\mathrm{pr}_{k-1}v) + B_k v_k \tag{3.13}$$

for all $v \in U_k$. Choose v such that $v_i = 0$ for all $i < k$ and $v_k \neq 0$. In this case,

$$E^i(v) = E^i(\mathrm{pr}_{i-1}v) = 0, \qquad \forall i \le k.$$

Hence, if we substitute such a control v in (3.12), using (3.10) and (3.13), we obtain the control law

$$u_k = B_k^t\lambda_{k+1}, \tag{3.14}$$

which we can plug in (3.1).

The next step is to obtain a discrete equation for λ_k. To do this we compare the multipliers λ_k and λ_{k+1} by subtracting (3.12) from the same formula with k shifted by -1. Using $v \in U_{k-1} \subset U_k$ as a test variation and formulas (3.10) and (3.13) we find that

$$\begin{aligned} &\lambda_{k+1}^t E^{k+1}(v) - \lambda_k^t E^k(v) = d_u\widetilde{\mathcal{J}}_s^k(v) - d_u\widetilde{\mathcal{J}}_s^{k-1}(v) \\ \iff &\big(\lambda_{k+1}^t P_k^{k+1} - \lambda_k^t + x_k^t Q_k\big)E^k(v) = 0. \end{aligned}$$

Since by Lemma 2 the endpoint map E^k is surjective for all k, the term in brackets must vanish.

Collecting everything gives equation (3.8). Taking into account that $E^{N+1}(u) = 0 = x_{N+1}$ and that we are assuming $x_0 = 0$, determines the boundary conditions.

To prove the converse it is enough to read all the relations backwards. By construction, any solution of equation (3.8) with the proper boundary conditions satisfies equation (3.11). Thus, using (3.14) as a definition for u we recover a control which is automatically a critical point for $\widetilde{\mathcal{J}}$.

Finally, we have to prove that this correspondence is bijective. From equation (3.1) and Assumption 1 it is clear that different controls give distinct trajectories and thus the correspondence is injective. Now let us prove the converse, clearly any solution of (3.15) gives a trajectory of (3.1) whose control is determined by (3.14). Assume that two solutions (x, λ) and $(x, \widetilde{\lambda})$ of equation (3.8) arise from the same control. Since (3.8) is linear, we can subtract the two solutions and obtain a new one which, now, has $x_k = 0$ for all k. This new solution has Lagrange multipliers $\mu_k := \lambda_k - \widetilde{\lambda}_k$, which satisfy

$$B_k B_k^t \mu_k = 0, \qquad \mu_k = (P_k^{k+1})^t \mu_{k+1}.$$

By Assumption 1 the matrix $B_0 B_0^t$ is invertible and hence $\mu_0 = 0$. Then again from Assumption 1 and the second equation it follows that $\mu_k = 0$ for all k. ∎

Under Assumption 1, we can rewrite system (3.8) as a forward equation. To do so, recall that $\Gamma_k = B_k B_k^t$, multiply the second equation by $(P_{k+1}^k)^t$ and plug in the new expression for λ_{k+1} into the first equation. This gives

$$\begin{cases} x_{k+1} = (P_k^{k+1} - \Gamma_k (P_{k+1}^k)^t Q_k) x_k + \Gamma_k (P_{k+1}^k)^t \lambda_k, \\ \lambda_{k+1} = (P_{k+1}^k)^t \lambda_k - (P_{k+1}^k)^t Q_k x_k. \end{cases} \tag{3.15}$$

We thus have

$$\begin{pmatrix} \lambda_{k+1} \\ x_{k+1} \end{pmatrix} = M_k \begin{pmatrix} \lambda_k \\ x_k \end{pmatrix},$$

where

$$M_k = \begin{pmatrix} (P_{k+1}^k)^t & 0 \\ \Gamma_k (P_{k+1}^k)^t & P_k^{k+1} \end{pmatrix} \begin{pmatrix} 1 & -Q_k \\ 0 & 1 \end{pmatrix}. \tag{3.16}$$

Both matrices in the product are symplectic, which makes M_k symplectic as well. Let us define the flow up to the point k as

$$\Phi_k = \begin{pmatrix} \Phi_k^1 & \Phi_k^2 \\ \Phi_k^3 & \Phi_k^4 \end{pmatrix} := \prod_{i=0}^{k-1} M_i. \tag{3.17}$$

Remark 6. We can reformulate Lemma 3 and the boundary value problem (3.8) in terms of Φ_k. We are looking for solutions of (3.8) with boundary values

$$x_0 = x_{k+1} = 0.$$

Writing as above Φ_{k+1} as a block matrix, shows that the existence of a solution is equivalent to the vanishing of the determinant of the block Φ^3_{k+1},

$$\ker \widetilde{\mathcal{J}}^k|_{\ker E^{k+1}} \neq 0 \iff \det \Phi^3_{k+1} = 0.$$

Remark 7. In the proofs of Theorem 1 and Theorem 2, we will need to deal with complex matrices. The proof of Lemma 3 extends to this case in a straightforward way since it only relies on the Lagrange multipliers rule, which still holds for maps from $\mathbb{C}^d$ to $\mathbb{C}^s$, and the non-degeneracy of the (Hermitian) scalar product. One needs only to replace transpose matrices with conjugate transpose ones.

4. A recursive formula for the index in LQR problems

The goal of this section is to prove the following result and deduce from it Theorems 1 and 2. Let Φ_k and M_k be the symplectic matrices defined in (3.17) and (3.16) respectively.

Theorem 3. *Let $\widetilde{\mathcal{J}}$ be as in (3.2). The negative inertia index of $\widetilde{\mathcal{J}}$ restricted to* $\ker E^{N+1}$ *is given by the following formula:*

$$\operatorname{ind}^- \widetilde{\mathcal{J}}|_{\ker E^{N+1}} = \sum_{k=1}^{N} \operatorname{ind}^- \mathcal{H}_k + \dim\big(\ker \Phi^3_k / (\ker \Phi^3_k \cap \ker \Phi^3_{k+1})\big),$$

where $\mathcal{H}_k$ is the quadratic form given by the matrix

$$\widetilde{\mathcal{H}}_k = \begin{cases} \big(\Gamma_0^\dagger + (P_1^2)^t \Gamma_1^\dagger P_1^2 - Q_1\big), & k = 1, \\ (\Phi^3_k)^t \big((P_k^{k+1})^t \Gamma_k^\dagger (P_k^{k+1}) - Q_k\big)\Phi^3_k + (\Phi^3_k)^t \Gamma^\dagger_{k-1} \Gamma_{k-1} \Phi^1_k, & k \neq 1, \end{cases}$$

restricted to the subspace $\ker(1 - B_k B_k^\dagger) P_k^{k+1} \Phi^3_k$.

As explained in Section 3.1, we can equivalently work on the space of controls, on $\ker E^{N+1}$, with the functional $\widetilde{\mathcal{J}}$ given in (3.2) or on the space of admissible trajectories (i.e., solutions of (3.1)) having $x_{N+1} = 0$ with the Jacobi matrix $\mathcal{I}$ defined by (3.3). We will choose the latter point of view.

The two main ingredients of Theorem 3 proof are the filtration corresponding to the U_k defined in (3.6) (seen as a filtration of the space on admissible curves $\mathcal{W}$) and Lemma 1.

Let us define the following family of subspaces:

$$\mathcal{W}_k = \{x \subseteq \mathbb{R}^{nN} : x \text{ solves (3.1) and } x_i = 0,\ \forall i \geq k+1\}.$$

We will apply iteratively Lemma 1 to the $\mathcal{W}_k$. It is straightforward to check that $\mathcal{W}_{k-1} \subset \mathcal{W}_k$ and $\bigcup_{k=1}^{N+1} \mathcal{W}_k = \mathcal{W}$. Thus the $\mathcal{W}_k$ provide a filtration of the space of admissible trajectories $\mathcal{W}$. Moreover, $\mathcal{W}_N$ corresponds to the space of admissible trajectories having $x_{N+1} = 0$.

Let us introduce some simplified notations as before. Let $\mathcal{I}^k$ be the principal $nk \times nk$ minor of $\mathcal{I}$ with coefficients given by (3.3). It corresponds to the functional $\tilde{\mathcal{J}}^k$ defined in (3.9) after the coordinate change. Set

$$\mathcal{Q}_k := \mathcal{I}^k|_{\mathcal{W}_k}, \qquad \mathcal{W}_k^\perp := \mathcal{W}_k^{\perp_{\mathcal{Q}_{k+1}}}, \qquad \mathcal{Q}_k^\perp := \mathcal{Q}_k|_{\mathcal{W}_{k-1}^\perp}.$$

Note that by definition $\mathcal{W}_k^\perp \subset \mathcal{W}_{k+1}$. Moreover, we have that

$$\mathcal{Q}_{k+1}|_{\mathcal{W}_k} = \mathcal{Q}_k.$$

Let us apply Lemma 1 with $\mathcal{Q} = \mathcal{Q}_{k+1}$ and $W = \mathcal{W}_k \subseteq \mathcal{W}_{k+1} = X$. We obtain via iteration

$$\begin{aligned}\operatorname{ind}^-\mathcal{I} = \operatorname{ind}^-\mathcal{Q}_1 + \sum_{k=2}^{N} \operatorname{ind}^-\mathcal{Q}_k^\perp \\ + \dim\big((\mathcal{W}_{k-1} \cap \mathcal{W}_{k-1}^\perp)/(\mathcal{W}_{k-1} \cap \mathcal{W}_{k-1}^\perp \cap \ker \mathcal{Q}_k)\big).\end{aligned}$$

What is left to do is to express every term using the discrete Hamiltonian system given in (3.8). First of all, let us describe the subspaces $\mathcal{W}_k^\perp \cap \mathcal{W}_k$ and $\mathcal{W}_k^\perp \cap \mathcal{W}_k \cap \ker \mathcal{Q}_{k+1}$. Let Π be the vertical subspace, i.e.,

$$\Pi = \{(\lambda, 0) \in \mathbb{R}^{2n} : \lambda \in \mathbb{R}^n\}.$$

Lemma 4. *We have the following identifications in terms of the Hamiltonian flow* (3.15)*:*

$$\begin{aligned}\mathcal{W}_k^\perp \cap \mathcal{W}_k &= \Phi_{k+1}(\Pi) \cap \Pi;\\ \mathcal{W}_k^\perp \cap \mathcal{W}_k \cap \ker(\mathcal{Q}_{k+1}) &= \Phi_{k+1}(\Pi) \cap M_{k+1}^{-1}(\Pi) \cap \Pi.\end{aligned}$$

In terms of the symplectic matrices in (3.17) *this reads*

$$\begin{aligned}\mathcal{W}_k^\perp \cap \mathcal{W}_k &= \ker(\mathcal{Q}_k) = \ker \Phi_{k+1}^3;\\ \mathcal{W}_k^\perp \cap \mathcal{W}_k \cap \ker(\mathcal{Q}_{k+1}) &= \ker \Phi_{k+2}^3 \cap \ker \Phi_{k+1}^3.\end{aligned}$$

Proof. A direct consequence of the definitions of $\mathcal{W}_k$ and $\mathcal{W}_k^\perp$ is that $\mathcal{W}_k^\perp \cap \mathcal{W}_k = \ker(\mathcal{Q}_k)$. Lemma 3 identifies the elements of the kernel with solutions of the equation (3.8) with $x_0 = x_{k+1} = 0$. In terms of the corresponding flow Φ_{k+1} we get

$$\ker(\mathcal{Q}_k) = \Phi_{k+1}(\Pi) \cap \Pi.$$

It follows that $\mathcal{W}_k^\perp \cap \mathcal{W}_k \cap \ker(\mathcal{Q}_{k+1}) = \ker(\mathcal{Q}_k) \cap \ker(\mathcal{Q}_{k+1})$ and the two equalities follow. ■

The next step is to identify $\mathcal{Q}_{k+1}^\perp$. The next Lemma provides a matrix representation involving the matrices Φ_k defined in (3.17).

Lemma 5. *Assume that $k > 0$. Let Φ_{k+1}^1 and Φ_{k+1}^3 be the upper and lower right minors of Φ_{k+1} as defined in (3.17). The negative index of $\mathcal{Q}_{k+1}$ restricted to $\mathcal{W}_k^\perp$ coincides with the negative index of the following quadratic form:*

$$\big\langle (\Phi_{k+1}^3)^t \big(((P_{k+1}^{k+2})^t \Gamma_{k+1}^\dagger P_{k+1}^{k+2} - Q_{k+1})\Phi_{k+1}^3 + \Gamma_k^\dagger \Gamma_k \Phi_{k+1}^1 \big)\lambda_0, \lambda_0 \big\rangle, \tag{4.1}$$

restricted to the kernel of $(1 - \Gamma_{k+1}^\dagger \Gamma_{k+1}) P_{k+1}^{k+2} \Phi_{k+1}^3$.

Proof. The first step of the proof is to identify the elements of $\mathcal{W}_k^\perp$. As one might expect, they are closely related to the symplectic matrix Φ_{k+1}. Recall that under Assumption 1, the endpoint map is surjective for any $k \geq 1$. This allows us to write down a three terms recurrence relation for $x \in \mathcal{W}_k^\perp \subset \mathcal{W}_{k+1}$. Indeed, testing against elements of $\mathcal{W}_k$, we find that $\langle x, \mathcal{I} y \rangle = 0$ for all $y \in \mathcal{W}_k$ if and only if

$$\begin{cases} S_1 x_1 + R_1 x_2 = 0, & \\ S_j x_j + R_{j-1}^* x_{j-1} + R_j x_{j+1} = 0, & \text{for } 1 < j \leq k, \\ x_j = 0, & \text{for } j > k+1, \\ (1 - \Gamma_{k+1}^\dagger \Gamma_{k+1}) P_{k+1}^{k+2} x_{k+1} = 0. & \end{cases}$$

Notice that the last equation of the system comes from the fact that x itself must be admissible. By Lemma 3 we can conclude that elements in $\mathcal{W}_k^\perp \subseteq \mathcal{W}_{k+1}$ are in one to one correspondence with trajectories of Φ_{k+1} having

$$x_0 = 0 \qquad \text{and} \qquad (1 - \Gamma_{k+1}^\dagger \Gamma_{k+1}) P_{k+1}^{k+2} x_{k+1} = 0,$$

i.e., $\mathcal{W}_k^\perp$ can be identified with a subspace of $\Phi_{k+1}(\Pi)$.

Let us compute now $\langle \mathcal{I} x, x \rangle$ for $x \in \mathcal{W}_k^\perp$. It is straightforward to notice that the only non-zero components of $\mathcal{I} x$ are the $k+1$-th and the $k+2$-th. It follows that

$$\langle \mathcal{I} x, x \rangle = \langle S_{k+1} x_{k+1} + R_k^t x_k, x_{k+1} \rangle.$$

By the discussion above, we can write x_k as a function of x_{k+1} multiplying by M_k^{-1} and obtain

$$\begin{pmatrix} \lambda_k \\ x_k \end{pmatrix} = \begin{pmatrix} (P_k^{k+1})^t \lambda_{k+1} + Q_k P_{k+1}^k (x_{k+1} - \Gamma_k \lambda_{k+1}) \\ P_{k+1}^k (x_{k+1} - \Gamma_k \lambda_{k+1}) \end{pmatrix}.$$

Finally, recalling the definition of S_k and R_k given in (3.3), we obtain that

$$\langle \mathcal{I}x, x \rangle = \big\langle \big((P_{k+1}^{k+2})^t \Gamma_{k+1}^\dagger P_{k+1}^{k+2} - Q_{k+1}\big) x_{k+1} + \Gamma_k^\dagger \Gamma_k \lambda_{k+1}, x_{k+1} \big\rangle. \tag{4.2}$$

■

It remains to compute $\operatorname{ind}^- \mathcal{Q}_1$. Clearly, using (3.3), if $x \in \mathcal{W}_1$

$$\langle \mathcal{I}x, x \rangle = \big\langle x_1, \big(\Gamma_0^\dagger + (P_1^2)^t \Gamma_1^\dagger P_1^2 - Q_1\big) x_1 \big\rangle \qquad \text{and} \qquad (1 - B_1 B_1^\dagger) P_1^2 x_1 = 0.$$

Now, each term in the formula for $\operatorname{ind}^- \mathcal{I}$ is identified. Collecting all the pieces together, finishes the proof of Theorem 3.

Remark 8. If we require that the matrices B_k are invertible for any k, we can easily rewrite the kernel in (4.1) in at least two *intrinsic* ways. In this case, $\Gamma_{k+1}^\dagger = \Gamma_{k+1}^{-1}$ and $\Gamma_k^\dagger \Gamma_k = 1$. The first way is related to classical discrete Sturm–Liouville theory and the number of sign changes of the fundamental solution of (3.15). Using the relation (3.17), we see that

$$\Phi_{k+2}^3 = P_{k+1}^{k+2} \Phi_{k+1}^3 + \Gamma_{k+1} (P_{k+2}^{k+1})^t (\Phi_{k+1}^1 - Q_{k+1} \Phi_{k+1}^3).$$

After collecting a common term $(\Phi_{k+1}^3)^t (P_{k+1}^{k+2})^t \Gamma_k^{-1}$ from the left in (4.1) we see that the kernel can be rewritten as

$$(\Phi_{k+1}^3)^t (P_{k+1}^{k+2})^t \Gamma_k^{-1} \Phi_{k+2}^3.$$

The second interpretation comes from symplectic geometry. The form can be interpreted as a *Maslov form*. This quadratic form is a certain invariant of triples of Lagrangian subspaces and measures the relative position of the third inside the Lagrange Grassmannian with respect to a reference frame given by the first two. Let Λ_1, Λ_2, Λ_3 be Lagrangian subspaces. Let us define $m(\Lambda_1, \Lambda_2, \Lambda_3)$ as a quadratic form on $(\Lambda_1 + \Lambda_3) \cap \Lambda_2$ as follows. If $p_2 = p_1 + p_3$ where $p_i \in \Lambda_i$, we set

$$m(\Lambda_1, \Lambda_2, \Lambda_3)(p_2) = \sigma(p_1, p_3).$$

It turns out that $\mathcal{Q}_{k+1}^{\perp}$ coincides with $m(\Phi_{k+1}(\Pi), M_{k+1}^{-1}(\Pi), \Pi)$. Indeed, the space $M_{k+1}^{-1}(\Pi) \cap (\Phi_{k+1}(\Pi) + \Pi)$ is given by.

$$\begin{aligned}
M_{k+1}^{-1}\begin{pmatrix}\lambda_{k+2}\\ 0\end{pmatrix} &= \begin{pmatrix}((P_{k+1}^{k+2})^t - Q_{k+1}P_{k+2}^{k+1}\Gamma_{k+1})\lambda_{k+2}\\ -P_{k+2}^{k+1}\Gamma_{k+1}\lambda_{k+2}\end{pmatrix}\\
&= \begin{pmatrix}\lambda_{k+1}+\nu\\ y_{k+1}\end{pmatrix} \in \Pi + \Phi_{k+1}(\Pi).
\end{aligned}$$

Thus the value of $m(\lambda_{k+2}) = \sigma((\lambda_{k+1}, y_{k+1}), (\nu, 0)) = -\langle \nu, y_{k+1}\rangle$ is determined by

$$\begin{aligned}
m(\lambda_{k+2}) &= -\big\langle \big((P_{k+1}^{k+2})^t - Q_{k+1}P_{k+2}^{k+1}\Gamma_{k+1}\big)\lambda_{k+2} - \lambda_{k+1}, y_{k+1}\big\rangle\\
&= \big\langle \big((P_{k+1}^{k+2})^t\Gamma_{k+1}^{-1}P_{k+1}^{k+2} - Q_{k+1}\big)y_{k+1}, y_{k+1}\big\rangle + \langle \lambda_{k+1}, x_{k+1}\rangle,
\end{aligned}$$

where we substituted λ_{k+2} using the relation $\lambda_{k+2} = -\Gamma_{k+1}^{-1}P_{k+1}^{k+2}y_{k+1}$ and obtained exactly (4.2).

4.1. Jacobi matrices via optimal control

In this section, we prove Theorem 1 and Theorem 2. The strategy is to choose a convenient optimal control problem and simplify all the formulas of the previous section.

Proof of Theorem 1. We prove the result for $\lambda = 0$. For $\lambda \neq 0$ it is enough to replace each S_i with $S_i - \lambda$.

Set $B_i = 1$ for all i. In this case, the are no restrictions on the domain of $\mathcal{H}_k$. By (3.3), we get the following relations:

$$\begin{aligned}
\Gamma_i &= 1,\\
R_i &= -(P_i^{i+1})^*,\\
Q_i &= 1 + R_iR_i^* - S_i.
\end{aligned}$$

We plug those formulas into (3.16) and simplify a little bit the expressions. Notice that, for $k > 1$, from equation (3.15) we get

$$\begin{aligned}
-R_kx_{k+1} = (P_k^{k+1})^t x_{k+1} &= \big((P_k^{k+1})^tP_k^{k+1} - Q_k\big)x_k + \lambda_k\\
&= \big(((P_k^{k+1})^tP_k^{k+1} - Q_k)\Phi_k^3 + \Phi_k^1\big)\lambda_0.
\end{aligned}$$

On the other hand, one easily sees that $-\langle x_k, R_kx_{k+1}\rangle$ coincides with $\langle \lambda_0, \mathcal{W}_k\lambda_0\rangle$. Hence,

$$\mathcal{H}_k = -(\Phi_k^3)^*R_k\Phi_{k+1}^3.$$

It only remains to prove that the formula above holds for $k = 1$. Since $\mathcal{I}|_{\mathcal{V}_1}$ coincides with S_1, the contribution of the first point is just $\operatorname{ind}^- S_1$. Extend our Jacobi matrix to include blocks S_0 and R_0, which can be any as long as $S_0^* = S_0$ and $\det(R_0) \neq 0$. Then, using the formulas for Φ_2 and Φ_1 from equation (1.2), we find that

$$-(\Phi_1^3)^* R_1 \Phi_2^3 = (R_0^{-1})^* S_1 (R_0^{-1})$$

and its index is equal to the one of S_1 as required by Theorem 3. ∎

Proof of Theorem 2. Again we prove just the case $\lambda = 0$. Recall that we are assuming R_k to be Hermitian. Take rows of B_k to be an orthonormal basis of $\operatorname{Im}(R_k)$ and $P_k^{k+1} = -R_k - L_k$ where $L_k R_k = 0$ and $L_k + R_k$ is invertible and Hermitian. These choices will have the following consequences:

- $L_k R_k = R_k L_k = 0$ and as a consequence $L_k^\dagger R_k = 0$;
- Γ_k is the orthogonal projection on $\operatorname{Im} R_k$. Thus to make this more evident we define $\pi_k = \Gamma_k$ and clearly $\pi_k^\dagger = \pi_k$;
- the two points imply that $\pi_k L_k = 0$;
- $\pi_k R_k = R_k \pi_k = R_k$;
- $(L_k + R_k)^{-1} = L_k^\dagger + R_k^\dagger$.

  Thanks to (3.3), we set

$$\begin{aligned} R_k + L_k &= -P_k^{k+1}, \\ Q_k &= \Gamma_{k-1} + R_k^2 - S_k. \end{aligned}$$

It is a straightforward check that the matrices M_k and Φ_{k+1} have the same form as in equation (1.4). Let us assume that $k > 1$, plugging the relations above inside (4.1), we find that

$$\widetilde{\mathcal{H}}_k = (\Phi_k^3)^* \big((S_k - \pi_{k-1})\Phi_k^3 + \pi_{k-1}\Phi_k^1\big),$$

which has to be restricted to the subspace $\ker(1-\pi_k)(R_k + L_k)\Phi_k^3$, i.e., to the space given by $\ker L_k \Phi_k^3$. Plugging in the expression for Q_k and P_k^{k+1} into (3.15) we find the following equation for x_{k+1}:

$$x_{k+1} = -R_k^\dagger(\lambda_k + (S_k - \pi_{k-1})x_k) - L_k x_k.$$

It follows that, if $L_k x_k = 0$ (and consequently $\pi_k x_k = x_k$) and $x_k = \Phi_k^3 \lambda_0$,

$$\begin{aligned} -\langle x_k, R_k x_{k+1}\rangle &= \langle x_k, \pi_k(\lambda_k + (S_k - \pi_{k-1})x_k)\rangle = \langle x_k, \lambda_k + (S_k - \pi_{k-1})x_k\rangle \\ &= \langle \lambda_0, \widetilde{\mathcal{H}}_k \lambda_0\rangle + \langle x_k, (1-\pi_{k-1})\lambda_k\rangle. \end{aligned}$$

So, for $k \geq 1$ the quadratic form $\mathcal{H}_k$ coincides with

$$-\langle x_k, R_k x_{k+1} + (1-\pi_{k-1})\lambda_k \rangle = -\langle \lambda_0, (\Phi_k^3)^*(R_k \Phi_{k+1}^3 + (1-\pi_{k-1})\Phi_k^1)\lambda_0 \rangle,$$

restricted to $\ker(1-\pi_k)\Phi_k^3$. It remains to check the case $k=1$. The contribution we are missing is the one of $\mathcal{I}|_{\mathcal{W}_1}$ wich coincides again with the index of S_1 restricted to $\{x_1 \in \mathbb{R}^n : (1-\pi_1)x_1 = 0\}$. Let us show that the formula for $\mathcal{H}_1$ still holds true, provided that B_0 is invertible (and thus $1-\pi_0 = 0$). Using (3.16), we have that $x_1 = \Gamma_0 (P_1^0)^* \lambda_0 = \Gamma_0 \lambda_1$ and using (3.3), we find $Q_1 = \Gamma_0^{-1} + R_1^2 - S_1^2$. So, if $x \in \mathcal{W}_1$, we can compute

$$-\langle x_1, R_1 x_2 \rangle = \langle x_1, \pi_1(\lambda_1 - (\Gamma_0^{-1} - S_1)x_1) \rangle = \langle x_1, S_1 x_1 \rangle = \langle x, \mathcal{I} x \rangle. \qquad \blacksquare$$

Acknowledgements. The authors would like to thanks Prof. Andrei Agrachev for the stimulating discussions and the anonymous reviewer for the careful reading of the first draft of the paper and for pointing out the connection with standard discrete Sturm–Liouville theory.

Funding. The work of the second author was supported through the CIDMA Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology (“FCT - Fundação para a Ciência e a Tecnologia") within the project UIDP/04106/2020 and UIDB/04106/2020.

References

[1] A. Agrachev, S. Baranzini, and I. Beschastnyi, Index theorems for graph-parametrized optimal control problems. *Nonlinearity* **36** (2023), no. 5, 2792–2838 Zbl 1515.49016 MR 4580663

[2] A. Agrachev and I. Beschastnyi, Jacobi fields in optimal control: Morse and Maslov indices. *Nonlinear Anal.* **214** (2022), article no. 112608 Zbl 1476.58011 MR 4322334

[3] S. Baranzini, Functional determinants for the second variation. 2022, arXiv:2212.13455

[4] R. L. Burden, J. D. Faires, and A. C. Reynolds, *Numerical analysis*. Prindle, Weber & Schmidt, Boston, MA, 1978 Zbl 0419.65001 MR 519124

[5] S. D. Burton, The determinant and volume of 2-bridge links and alternating 3-braids. *New York J. Math.* **24** (2018), 293–316 Zbl 1392.57004 MR 3778507

[6] S. T. Galtung and K. Grunert, A numerical study of variational discretizations of the Camassa–Holm equation. *BIT* **61** (2021), no. 4, 1271–1309 Zbl 1477.35212 MR 4332137

[7] A. Hucht, The square lattice Ising model on the rectangle III: Hankel and Toeplitz determinants. *J. Phys. A* **54** (2021), no. 37, article no. 375201 Zbl 1520.82012 MR 4318580

[8] J. H. Maddocks, Restricted quadratic forms, inertia theorems, and the Schur complement. *Linear Algebra Appl.* **108** (1988), 1–36 Zbl 0653.15020 MR 959695

[9] Y. Murata, *Optimal control methods for linear discrete-time economic systems*. Springer, New York, 1982 Zbl 0494.90001 MR 784615

[10] H. Schulz-Baldes, Rotation numbers for Jacobi matrices with matrix entries. *Math. Phys. Electron. J.* **13** (2007), article no. 5 Zbl 1141.15024 MR 2366118

[11] H. Schulz-Baldes, Sturm intersection theory for periodic Jacobi matrices and linear Hamiltonian systems. *Linear Algebra Appl.* **436** (2012), no. 3, 498–515 Zbl 1232.15010 MR 2854887

[12] H. Schulz-Baldes and L. Urban, Space versus energy oscillations of Prüfer phases for matrix Sturm–Liouville and Jacobi operators. *Electron. J. Differential Equations* (2020), article no. 76 Zbl 1452.34037 MR 4126540

Received 20 February 2023; revised 4 December 2023.

Stefano Baranzini
Dipartimento di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, Italy; stefano.baranzini@unito.it

Ivan Beschastnyi
Departamento de Matemática, Universidade de Aveiro, Campus Santiago, 3810-193 Aveiro, Portugal; ibeschastnyi@ua.pt