ESAIM: COCV 18 (2012) 501–519
DOI: 10.1051/cocv/2011104

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

RAYLEIGH PRINCIPLE FOR LINEAR HAMILTONIAN SYSTEMS WITHOUT CONTROLLABILITY*

Werner Kratz[1] and Roman Šimon Hilscher[2]

Abstract. In this paper we consider linear Hamiltonian differential systems without the controllability (or normality) assumption. We prove the Rayleigh principle for these systems with Dirichlet boundary conditions, which provides a variational characterization of the finite eigenvalues of the associated self-adjoint eigenvalue problem. This result generalizes the traditional Rayleigh principle to possibly abnormal linear Hamiltonian systems. The main tools are the extended Picone formula, which is proven here for this general setting, results on piecewise constant kernels for conjoined bases of the Hamiltonian system, and the oscillation theorem relating the number of proper focal points of conjoined bases with the number of finite eigenvalues. As applications we obtain the expansion theorem in the space of admissible functions without controllability and a result on coercivity of the corresponding quadratic functional.

Mathematics Subject Classification. 34L05, 34C10, 34L10, 34B09.

Received October 25, 2010.
Published online July 22, 2011.

1. Introduction and main result

This paper deals with continuous time linear Hamiltonian systems and related quadratic functionals and eigenvalue problems with Dirichlet boundary conditions. The main result is the Rayleigh principle describing a variational method for finding the eigenvalues. A key feature of this paper resides in the fact that we *do not* assume controllability (or equivalently normality) of the Hamiltonian system, as opposed to the existing literature on this subject such as in [8], Theorem 1, or [7], Theorem 7.7.1. Our main result (Thm. 1.1) is based on several recent topics from the theory of linear Hamiltonian systems without normality. In particular, these topics are:

(i) a new extended global Picone formula involving the finite eigenfunctions of the associated eigenvalue problem (Thm. 3.1),

Keywords and phrases. Linear Hamiltonian system, Rayleigh principle, self-adjoint eigenvalue problem, proper focal point, conjoined basis, finite eigenvalue, oscillation theorem, controllability, normality, quadratic functional.

* *Research supported by the Deutsche Forschungsgemeinschaft (DFG) under grant KR 1157/2–1, by the Czech Science Foundation under grant 201/09/J009, and by the research project MSM 0021622409 of the Ministry of Education, Youth, and Sports of the Czech Republic.*

[1] Department of Applied Analysis, Faculty of Mathematics and Economics, University of Ulm, 89069 Ulm, Germany.
werner.kratz@uni-ulm.de
[2] Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic. hilscher@math.muni.cz

Article published by EDP Sciences
© EDP Sciences, SMAI 2011

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

(ii) the oscillation theorem relating the number of proper focal points of conjoined bases of the Hamiltonian system with the number of finite eigenvalues of the eigenvalue problem (Thm. 2.9),
(iii) the geometric characterization of finite eigenvalues in terms of finite eigenfunctions (Thm. A.2 in the Appendix),
(iv) the positivity of quadratic functionals (Prop. 2.2).

As applications of our main result we derive the expansion theorem in the space of admissible functions (Thm. 4.3) and the equivalence between the positivity and coercivity of a quadratic functional (Thm. 4.5).

Let us introduce the subject of this paper in more detail. Let be given a dimension $n \in \mathbb{N}$, a fixed interval $[a,b]$ with $a < b$, and

$$\text{real } n \times n \text{ matrices } A, B, C \in \mathrm{C_p} \text{ on } [a,b] \text{ such that } B(t) \text{ and } C(t) \text{ are symmetric,} \tag{1.1}$$

where $\mathrm{C_p}$ is the set of piecewise continuous functions. Given (1.1), we consider the *linear Hamiltonian system*

$$x' = A(t)\,x + B(t)\,u, \quad u' = C(t)\,x - A^T(t)\,u, \quad t \in [a,b], \tag{H}$$

for real n-vector-valued functions $x, u \in \mathrm{C_p^1}$ (piecewise continuously differentiable functions), and the quadratic functional

$$\mathcal{F}_0(z) := \int_a^b \{x^T C x + u^T B u\}\,(t)\,\mathrm{d}t$$

for *admissible* pairs $z = (x,u)$, *i.e.*, the functions $x \in \mathrm{C_p^1}$, $u \in \mathrm{C_p}$ satisfy the equation of motion $x'(t) = A(t)\,x(t) + B(t)\,u(t)$ on $[a,b]$. Using the concept of piecewise continuous functions implies that we adopt in this paper the following common agreement concerning the points of discontinuity. A function $f \in \mathrm{C_p}$ is evaluated at a point t at which it is not continuous by the one sided limits $f(t^+)$ and $f(t^-)$.

The oscillation theory of system (H) is very well understood, see *e.g.* [2,7,8,11,12], when the system (H) satisfies the *(complete) controllability* condition: If $(x(\cdot) \equiv 0, u)$ is a solution of the system (H) on a nondegenerate subinterval of $[a,b]$, then also $u(\cdot) \equiv 0$. This is also called the *(identical) normality* of the system (H). We stress that by an "abnormal" system we mean a system which is not normal. As it is shown in [7], Theorem 4.1.3, if the Legendre condition

$$B(t) \geq 0, \textit{ i.e., } B(t) \text{ is symmetric and nonnegative definite, for all } t \in [a,b] \tag{1.2}$$

holds, then the above normality condition is equivalent to the fact that conjoined bases of the system (H), *i.e.*, the matrix solutions (X,U) which have X^TU symmetric and $\operatorname{rank}(X^T\,U^T) = n$, have $X(t)$ invertible everywhere in $[a,b]$ except possibly at isolated points $t_0 \in [a,b]$. Such an isolated point t_0 where $X(t_0)$ is singular is then called a *focal point* of (X,U) and the defect of $X(t_0)$, $\operatorname{def} X(t_0) := \dim \operatorname{Ker} X(t_0)$, is its multiplicity. When the normality condition is removed, then $X(\cdot)$ may be singular on an interval or even throughout $[a,b]$, which invalidates the oscillation theory based on the above focal point notion.

In [9], Theorem 3 (see Thm. 2.1), it is shown that under the Legendre condition (1.2) the kernel of $X(\cdot)$ is *piecewise constant* on $[a,b]$ and that the Moore–Penrose generalized inverse $X^\dagger(\cdot)$ of $X(\cdot)$ can be used to develop the theory. In particular, the matrix function $X^\dagger(\cdot)$ is differentiable on intervals where the kernel of $X(\cdot)$ is constant (see [9], Lem. 6). This leads to the notion of no (generalized or proper) focal points in $(a,b]$ introduced in [9] as

$$\operatorname{Ker} X(t) \subseteq \operatorname{Ker} X(\tau) \quad \text{for all } t, \tau \in [a,b],\ \tau \leq t, \tag{1.3}$$

to which we refer to as the "kernel condition" and which characterizes the positivity of $\mathcal{F}_0$ (see Prop. 2.2). Moreover, the nonnegativity of $\mathcal{F}_0$ was characterized in [9] in terms of

$$x(t) \in \operatorname{Im} X(t) \quad \text{for all } t \in [a,b] \tag{1.4}$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

for every admissible $z = (x, u)$ with $x(a) = 0 = x(b)$. We refer to (1.4) as the "image condition" (see Prop. 2.3). Other recent results related to the theory of linear Hamiltonian systems without controllability, such as the Riccati matrix differential equations, can be found in [5], Section 6.

The notion of *multiplicities* of proper focal points for the abnormal case was introduced in [14,15]. The associated self-adjoint eigenvalue problem

$$(\mathrm{H}_\lambda), \quad x(a) = 0 = x(b), \tag{E}$$

where (H_λ) is the linear Hamiltonian system

$$x' = A(t)\,x + B(t)\,u, \quad u' = C(t)\,x - A^T(t)\,u - \lambda W(t)\,x, \quad t \in [a, b], \tag{H_λ}$$

was first studied in [14,15] and later in [10] in a more general setting on time scales. In these references, the main result is the so-called oscillation theorem, which says that the number of proper focal points in $(a, b]$ of a special conjoined basis of the system (H_λ) is the same as the number of finite eigenvalues of (E) which are less or equal to λ (see Thm. 2.9). These results require that $W(\cdot)$ is a real and symmetric $n \times n$ matrix-function on $[a, b]$ such that $W \in \mathrm{C_p}$ and

$$W(t) \geq 0 \quad \text{for all } t \in [a, b]. \tag{1.5}$$

For the precise definition of finite eigenvalues and finite eigenfunctions of (E) we refer to Section 2.2. We always count the finite eigenvalues as well as the proper focal points including their multiplicities.

The *Rayleigh principle* presented in this paper is a natural continuation of the above project for possibly abnormal linear Hamiltonian systems. We consider the quadratic functional

$$\mathcal{F}_\lambda(z) := \mathcal{F}_0(z) - \lambda\,\langle z, z\rangle_W, \quad \langle z, \tilde{z}\rangle_W := \int_a^b \{x^T W \tilde{x}\}\,(t)\,\mathrm{d}t, \tag{1.6}$$

where $z = (x, u)$ and $\tilde{z} = (\tilde{x}, \tilde{u})$ are admissible and $\lambda \in \mathbb{R}$. We say that two functions z and $\tilde{z}$ are orthogonal (with respect to the bilinear form $\langle\cdot,\cdot\rangle_W$) and write $z \perp \tilde{z}$, provided $\langle z, \tilde{z}\rangle_W = 0$. If we denote by

$$\mathcal{A} := \big\{z, \text{ such that } z = (x, u) \text{ is admissible and } x(a) = 0 = x(b)\big\} \tag{1.7}$$

the space of admissible functions for the functional $\mathcal{F}_\lambda$ (*i.e.*, admissible with Dirichlet boundary conditions), then the main result of this paper reads as follows.

Theorem 1.1 (Rayleigh principle)**.** *Assume* (1.5) *and suppose that the functional $\mathcal{F}_{\underline{\lambda}}$ is positive definite for some $\underline{\lambda} < 0$,* i.e.*, $\mathcal{F}_{\underline{\lambda}}(z) > 0$ for all $z \in \mathcal{A}$ with $x(\cdot) \not\equiv 0$. Let $\lambda_1 \leq \ldots \leq \lambda_m \leq \ldots$ be the finite eigenvalues of the eigenvalue problem* (E) *with the corresponding orthonormal finite eigenfunctions $z_1, \ldots, z_m, \ldots$ Then for each $m \in \mathbb{N} \cup \{0\}$*

$$\lambda_{m+1} = \min\left\{\frac{\mathcal{F}_0(z)}{\langle z, z\rangle_W}, \quad z \in \mathcal{A},\ (Wx)(\cdot) \not\equiv 0,\ \textit{and}\ z \perp z_1, \ldots, z_m\right\}. \tag{1.8}$$

When $m = 0$, the orthogonality condition in (1.8) becomes empty. The number of finite eigenvalues may be finite (or even equal to zero) by our general assumptions, see Corollary 4.1 below. The assumption on the positivity of $\mathcal{F}_{\underline{\lambda}}$ allows one to apply the oscillation theorem and ensures that the finite eigenvalues of (E) are isolated and bounded from below (see Thm. 2.9), so that the list $\lambda_1 \leq \ldots \leq \lambda_m \leq \ldots$ makes sense.

The above theorem is a direct generalization of [8], Theorem 1, or [7], Theorem 7.7.1 (reduced to the Dirichlet boundary conditions), to linear Hamiltonian systems without the normality assumption. Our Theorem 1.1 can also be viewed as the continuous time counterpart of the Rayleigh principle for discrete symplectic systems in [1], Theorem 4.6. And as in the latter reference, the Rayleigh principle is a promising tool for a general Sturmian theory for possibly abnormal linear Hamiltonian systems.

The set up of the paper is the following. In the next section we recall the central notions of proper focal points for conjoined bases of (H) and finite eigenvalues of (E) from [15], and we present several other needed auxiliary

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

results from this reference and from [9]. In Section 3 we establish an extended Picone formula, which involves the finite eigenfunctions of (E) and which is the key tool for the proof of the Rayleigh principle. This proof is then presented in Section 4, where we also provide applications of the Rayleigh principle in the form of the expansion theorem and a result on the coercivity of $\mathcal{F}_0$. In order to keep this paper complete and self-contained, we include in Appendix A the proof of the geometric characterization of finite eigenvalues used in the results of this paper, and in Appendix B the proof of the oscillation theorem (Thm. 2.9). These proofs are simpler and shorter than those in [15] and they are essentially extracted from more general oscillation results on time scales in [10], which we specialize to the continuous time setting.

2. Basic notions and auxiliary tools

In this section we present the basic terminology and other auxiliary topics needed in this paper.

2.1. Proper focal points

The following is a fundamental result behind the general proper focal point definition.

Theorem 2.1 (piecewise constant kernel)**.** *Assume* (1.2). *Then every conjoined basis* (X,U) *of* (H) *has* $\operatorname{Ker} X(\cdot)$ *piecewise constant on* $[a,b]$. *More precisely, there are points* $a=t_0<t_1<\cdots<t_k=b$ *such that*

$$\operatorname{Ker} X(t)\equiv\operatorname{Ker} X(t_i^-)\subseteq\operatorname{Ker} X(t_i)\quad \text{for all } t\in(t_{i-1},t_i),\ i=1,\dots,k, \tag{2.1}$$

$$\operatorname{Ker} X(t)\equiv\operatorname{Ker} X(t_i^+)\subseteq\operatorname{Ker} X(t_i)\quad \text{for all } t\in(t_i,t_{i+1}),\ i=0,\dots,k-1. \tag{2.2}$$

Proof. See [9], Theorem 3. □

From the above result it is clear that the Legendre condition (1.2) is essential for this theory and it will be assumed further on. According to [15], Definition 1.1, a point $t_0\in(a,b]$ is called a *proper focal point* of a conjoined basis (X,U) of (H), if $\operatorname{Ker} X(t_0^-)\subsetneqq\operatorname{Ker} X(t_0)$, and then $m:=\operatorname{def} X(t_0)-\operatorname{def} X(t_0^-)$ is its multiplicity. This definition means that the multiplicity of $t_0\in(a,b]$ as a proper focal point of (X,U) is "somehow" the dimension of vectors which are in the kernel of $X(t_0)$ but which are not in the kernel of $X(t_0^-)$. More precisely,

$$m=\dim\big([\operatorname{Ker} X(t_0^-)]^\perp\cap\operatorname{Ker} X(t_0)\big). \tag{2.3}$$

When a conjoined basis does not have proper focal points in $(a,b]$, then by (2.1), $\operatorname{Ker} X(t^-)=\operatorname{Ker} X(t)$ for all $t\in(a,b]$, which together with the inclusion in (2.2) yields the kernel condition (1.3). And this kernel condition characterizes the positivity of $\mathcal{F}_0$.

We say that the functional $\mathcal{F}_0$ is positive definite (and write $\mathcal{F}_0>0$) if $\mathcal{F}_0(z)>0$ for every $z\in\mathcal{A}$ with $x(\cdot)\not\equiv 0$. We say that $\mathcal{F}_0$ is nonnegative (and write $\mathcal{F}_0\geq 0$) if $\mathcal{F}_0(z)\geq 0$ for all $z\in\mathcal{A}$.

Proposition 2.2 (positivity)**.** *We have* $\mathcal{F}_0>0$ *if and only if condition* (1.2) *holds and there exists a conjoined basis of* (H) *which has no proper focal points in* $(a,b]$.

Proof. See [9], Theorem 1 and Remark 3 (i). □

A similar result to Proposition 2.2 holds for the nonnegativity of $\mathcal{F}_0$, but with the image condition (1.4) instead of the kernel condition (1.4).

Proposition 2.3 (nonegativity)**.** *We have* $\mathcal{F}_0\geq 0$ *if and only if condition* (1.2) *holds and there exists a conjoined basis of* (H) *satisfying the image condition* (1.4) *for every admissible* $z=(x,u)$ *with* $x(a)=0=x(b)$.

Proof. See [9], Theorem 2, and the continuous time version of [13], Corollary 4.3. □

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

In fact, the conjoined basis in Propositions 2.2 and 2.3 can be always chosen to be the *principal solution* $(\hat{X}, \hat{U})$ of the system (H), which is the solution of (H) given by the initial conditions $\hat{X}(a) = 0$ and $\hat{U}(a) = I$. To the contrary with the traditional theory (see *e.g.* [7], Thm. 8.2.6), in the general abnormal setting the nonnegativity of $\mathcal{F}_0$ is not equivalent to the nonexistence of proper focal points in the open interval (a, b), see for example [9], Remark 12, where $[a, b] = [0, 2\pi]$, $\mathcal{F}_0 \geq 0$, but $t_0 = \pi \in (0, 2\pi)$ is a proper focal point of $(\hat{X}, \hat{U})$. However, the sufficiency of those conditions is clear from the global Picone formula (Thm. 2.11) below and from the relation between the kernel condition and the image condition (Lem. 2.5).

Corollary 2.4. *Assume that* (1.2) *holds and there exists a conjoined basis of* (H) *which has no proper focal points in* (a, b). *Then* $\mathcal{F}_0 \geq 0$.

Next we establish the relation between the kernel condition (1.3) and the image condition (1.4).

Lemma 2.5. *Let* (X, U) *be a conjoined basis of* (H) *satisfying the kernel condition* (1.3). *Then for any admissible* $z = (x, u)$ *with* $x(a) \in \operatorname{Im} X(a)$ *we have* $x(t) \in \operatorname{Im} X(t)$ *for all* $t \in [a, b]$.

Proof. This follows from [9], Corollary 4. From this reference we have that for every $t_0 \in [a, b]$ the reachable set $E_a(t_0)$ at the point t_0 is equal to $\operatorname{Im} X(t_0)$, where $E_a(t_0)$ is by definition the set of vectors $d \in \mathbb{R}^n$ for which there exists an admissible $z = (x, u)$ with $x(a) \in \operatorname{Im} X(a)$ and $x(t_0) = d$. □

2.2. **Finite eigenvalues**

Next we proceed with the properties of the eigenvalue problem (E) introduced in Section 1. Here we use the algebraic definition of finite eigenvalues from [10], Definition 2.4, as opposed to the geometric definition in [15], Definition 1.2. However, both definitions are equivalent under condition (1.5), as it is shown in Theorem A.2 in the appendix or in [10], Theorem 5.2. Let $\big(\hat{X}(\cdot, \lambda), \hat{U}(\cdot, \lambda)\big)$ be the principal solution of (H_λ), *i.e.*, $\hat{X}(a, \lambda) = 0$ and $\hat{U}(a, \lambda) = I$ for all $\lambda \in \mathbb{R}$.

Definition 2.6. A number $\lambda_0 \in \mathbb{R}$ is called a *finite eigenvalue* of the eigenvalue problem (E) provided

$$\theta(\lambda_0) := r(b) - \operatorname{rank} \hat{X}(b, \lambda_0) > 0, \qquad \text{where} \qquad r(b) := \max_{\lambda \in \mathbb{R}} \operatorname{rank} \hat{X}(b, \lambda). \tag{2.4}$$

In this case we call $\theta(\lambda_0)$ the *algebraic multiplicity* of the finite eigenvalue λ_0.

The above definition is motivated by the discrete time theory in [3], Definition 2. A solution $z(\cdot, \lambda_0) = (x(\cdot, \lambda_0), u(\cdot, \lambda_0))$ of (E) with $\lambda = \lambda_0$ is called a *finite eigenfunction* corresponding to the finite eigenvalue λ_0 provided $W(\cdot)\, x(\cdot, \lambda_0) \not\equiv 0$, and then the dimension of

$$\big\{\, W(\cdot)\, x(\cdot, \lambda_0), \quad z = (x, u) \text{ solves (E) with } \lambda = \lambda_0 \,\big\} \tag{2.5}$$

is called the *geometric multiplicity* of λ_0.

Remark 2.7. (i) Under condition (1.5), the eigenvalue problem (E) enjoys traditional properties of self-adjoint differential eigenvalue problems, such as that all finite eigenvalues are real, the algebraic and geometric multiplicities of a finite eigenvalue λ_0 are equal, and the finite eigenfunctions corresponding to different finite eigenvalues are orthogonal with respect to the bilinear form $\langle \cdot, \cdot \rangle_W$ defined in (1.6), see [15], Remark 1.3, and Appendix A. In particular, for every finite eigenvalue λ_0 with multiplicity $\theta(\lambda_0) \geq 1$ there are exactly $\theta(\lambda_0)$ linearly independent finite eigenfunctions, which can be orthonormalized by the standard procedure.

(ii) As it is proven in [10], Corollary 5.5, under the normality assumption our Definition 2.6 reduces to the classical definition of eigenvalues of (E) *e.g.* in [7], pp. 42–43, where $r(b) = n$ and λ_0 is an eigenvalue of (E) if and only if the matrix $\hat{X}(b, \lambda_0)$ is singular.

Remark 2.8. The number of finite eigenvalues of (E) depends in general on $W(\cdot)$ but also on $B(\cdot)$. For example, if $B(\cdot) \equiv 0$, then every admissible $z = (x, u)$ with $x(a) = 0$ has $x(\cdot) \equiv 0$, so that there are no finite eigenvalues at all in this case.

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

2.3. **Oscillation theorem**

The following oscillation theorem is an important tool for our theory. Let $\big(\hat{X}(\cdot,\lambda),\hat{U}(\cdot,\lambda)\big)$ be the principal solution of (H$_\lambda$). We will always count the proper focal points of conjoined bases of (H$_\lambda$) as well as the finite eigenvalues of (E) including their multiplicities. Denote by

$$n_1(\lambda) := \text{the number of proper focal points of } \big(\hat{X}(\cdot,\lambda),\hat{U}(\cdot,\lambda)\big) \text{ in } (a,b], \tag{2.6}$$

$$n_2(\lambda) := \text{the number of finite eigenvalues of (E) which are less or equal to } \lambda. \tag{2.7}$$

Theorem 2.9 (oscillation theorem)**.** *Assume* (1.2) *and* (1.5). *Then*

$$n_1(\lambda) = n_2(\lambda) \quad \text{for all } \lambda \in \mathbb{R} \tag{2.8}$$

if and only if there exists $\underline{\lambda} < 0$ *such that the functional* $\mathcal{F}_{\underline{\lambda}}$ *is positive definite. In this case the finite eigenvalues of* (E) *are bounded from below.*

Proof. See Appendix B, where we provide the proof of this oscillation theorem (or [15], Cor. 1.7, or [10], Cor. 6.4). □

The crucial assumption $\mathcal{F}_{\underline{\lambda}} > 0$ for some $\underline{\lambda} < 0$ in Theorem 1.1 (and in Thm. 2.9) is satisfied if $W(t) > 0$ on $[a,b]$. More precisely, we have the following lemma (compare with [10], Thm. 9.5).

Lemma 2.10 (positivity)**.** *Assume* (1.2) *and* $W(t) > 0$ *on* $[a,b]$. *Then there exists* $\omega > 0$ *and* $\underline{\lambda} < 0$ *such that for all* $\lambda \le \underline{\lambda}$

$$\mathcal{F}_\lambda(z) \ge \omega\,(-\lambda)\int_a^b |x(t)|^2\,\mathrm{d}t$$

for all $z = (x,u) \in \mathrm{C_p}$. *In particular,* $\mathcal{F}_\lambda(z) > 0$ *if* $x(\cdot) \not\equiv 0$.

Proof. By continuity, there exists $\omega > 0$ such that $W(t) \ge 2\omega I$ on $[a,b]$. Hence, by (1.2), for a piecewise continuous $z = (x,u)$ we have that

$$\mathcal{F}_\lambda(z) \ge \int_a^b \{x^T C x\}\,(t)\,\mathrm{d}t - \lambda\int_a^b 2\omega\,|x(t)|^2\,\mathrm{d}t \ge \omega\,(-\lambda)\int_a^b |x(t)|^2\,\mathrm{d}t,$$

provided $\lambda \le \underline{\lambda}$ for a sufficiently small $\underline{\lambda} < 0$. □

2.4. **Other auxiliary tools**

Next we present other important tools which we need in the proof of the Rayleigh principle. The first result of this kind is a global Picone formula providing a lower bound (the nonnegativity) for the functional $\mathcal{F}_0$. For a function f on $[a,b]$ we use the notation $f(t)\,|_a^b = f(b) - f(a)$.

Theorem 2.11 (global Picone formula)**.** *Assume* (1.2). *Let* (X,U) *be a conjoined basis of* (H) *and let* $z = (x,u)$ *be admissible satisfying the image condition* (1.4). *Then*

$$\mathcal{F}_0(z) \ge \int_a^b \{w^T B w\}\,(t)\,\mathrm{d}t + (x^T U X^\dagger x)\,(t)\,|_a^b,$$

where $w := u - UX^\dagger x$ *on* $[a,b]$. *If, in addition,* (X,U) *satisfies the kernel condition* (1.3), $\int_a^b \{w^T B w\}\,(t)\,\mathrm{d}t = 0$, *and* $x(b) = 0$, *then* $x(t) \equiv 0$ *on* $[a,b]$.

Proof. This global Picone formula is a special case of the corresponding time scale result in [13], Theorem 3.19, which is here reduced to the continuous time setting. The idea of the proof can also be followed in [9], Proposition 6, where the special choice of the conjoined basis $(X,U) = (\hat{X},\hat{U})$ is considered. □

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

For $z = (x, u)$ and $\hat{z} = (\hat{x}, \hat{u})$ we define the function

$$\Lambda(z, \hat{z})(t) := \{x^T C\hat{x} + u^T B\hat{u}\}(t), \quad t \in [a, b]. \tag{2.9}$$

Lemma 2.12. *Let $z = (x, u)$ be admissible and suppose that $\hat{z} = (\hat{x}, \hat{u}) \in \mathrm{C}^1_{\mathrm{p}}$. Then*

$$\int_a^b \Lambda(z, \hat{z})(t)\,\mathrm{d}t = (x^T\hat{u})(t)\Big|_a^b - \int_a^b \{x^T(\hat{u}' - C\hat{x} + A^T\hat{u})\}(t)\,\mathrm{d}t. \tag{2.10}$$

Proof. Identity (2.10) follows by the integration by parts formula. □

Next we derive formulas for the values of $\mathcal{F}_0(z)$ and $\langle z, z\rangle_W$ when z is a linear combination of finite eigenfunctions of (E).

Lemma 2.13. *Let $z_1, \dots, z_m$ be orthonormal finite eigenfunctions of (E) corresponding to the (not necessarily distinct and not necessarily consecutive) finite eigenvalues $\lambda_1, \dots, \lambda_m$. For any $\beta_1, \dots, \beta_m \in \mathbb{R}$ we set $\hat{z} := \sum_{i=1}^m \beta_i z_i$. Then $\hat{z} = (\hat{x}, \hat{u})$ is admissible, $\hat{x}(a) = 0 = \hat{x}(b)$, and*

$$\mathcal{F}_0(\hat{z}) = \sum_{i=1}^m \lambda_i \beta_i^2 \quad \text{and} \quad \langle \hat{z}, \hat{z}\rangle_W = \sum_{i=1}^m \beta_i^2. \tag{2.11}$$

Proof. Both formulas follow from the orthonormality relation between the finite eigenfunctions, *i.e.*, $\langle z_i, z_j\rangle_W = \delta_{ij}$. More precisely, the first formula in (2.11) is a consequence of Lemma 2.12, because

$$\begin{aligned}\mathcal{F}_0(\hat{z}) &= (\hat{x}^T\hat{u})(t)\Big|_a^b - \int_a^b \{\hat{x}^T(\hat{u}' - C\hat{x} + A^T\hat{u})\}(t)\,\mathrm{d}t \\ &= -\int_a^b \sum_{i=1}^m \beta_i \{\hat{x}^T(u_i' - Cx_i + A^T u_i)\}(t)\,\mathrm{d}t \overset{(\mathrm{H}_{\lambda_i})}{=} \sum_{i=1}^m \beta_i \lambda_i \langle \hat{z}, z_i\rangle_W = \sum_{i=1}^m \lambda_i \beta_i^2.\end{aligned}$$

The second formula in (2.11) is straightforward. □

3. Extended global Picone formula

In this section we establish an extended global Picone formula, which involves the finite eigenfunctions of (E), comparing to the global Picone formula in Theorem 2.11. For the Dirichlet boundary conditions it is a generalization of [7], Theorem 2.2.3, to the case of possibly noninvertible $X(\cdot)$ on $[a, b]$.

Theorem 3.1 (extended global Picone formula)**.** *Assume (1.2) and (1.5). Let (X, U) be a conjoined basis of (H_λ) for some fixed $\lambda \in \mathbb{R}$. Let $\lambda_1 \le \cdots \le \lambda_m$ be finite eigenvalues of (E) with the corresponding orthonormal finite eigenfunctions $z_1, \dots, z_m$, that is, for every $i \in \{1, \dots, m\}$*

$$x_i' = Ax_i + Bu_i, \quad u_i' = Cx_i - A^T u_i - \lambda_i W x_i \quad \text{on } [a, b], \tag{3.1}$$

$$x_i(a) = 0 = x_i(b), \quad (Wx_i)(\cdot) \not\equiv 0. \tag{3.2}$$

For any $\beta_1, \dots, \beta_m \in \mathbb{R}$ we set $\hat{z} := \sum_{i=1}^m \beta_i z_i$. Finally, let $z = (x, u)$ be admissible with $x(a) = 0 = x(b)$ and $z \perp z_1, \dots, z_m$ and such that $\tilde{z} = (\tilde{x}, \tilde{u}) := z + \hat{z}$ satisfies the image condition

$$\tilde{x}(t) \in \operatorname{Im} X(t) \quad \text{for all } t \in [a, b]. \tag{3.3}$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Then we have the inequality

$$\mathcal{F}_\lambda(z) \geq \int_a^b \{\tilde{w}^T B \tilde{w}\}(t)\,\mathrm{d}t + \sum_{i=1}^m (\lambda - \lambda_i)\,\beta_i^2, \tag{3.4}$$

where $\tilde{w} := \tilde{u} - UX^\dagger \tilde{x}$ *on* $[a,b]$.

Proof. Let $z = (x,u)$ be admissible with $x(a) = 0 = x(b)$ and $z \perp z_1, \ldots, z_m$. First note that $\tilde{z} = z + \hat{z}$ is admissible and, by (3.2), $\tilde{x}(a) = x(a) + \hat{x}(a) = 0$ and $\tilde{x}(b) = x(b) + \hat{x}(b) = 0$. By the global Picone formula (Thm. 2.11) applied to (X,U), system (H_λ), and the admissible $\tilde{z}$ (using (3.3)), we have

$$\mathcal{F}_\lambda(\tilde{z}) \geq \int_a^b \{\tilde{w}^T B \tilde{w}\}(t)\,\mathrm{d}t + (\tilde{x}^T U X^\dagger \tilde{x})(t)\,\big|_a^b = \int_a^b \{\tilde{w}^T B \tilde{w}\}(t)\,\mathrm{d}t, \tag{3.5}$$

where $\tilde{w} := \tilde{u} - UX^\dagger \tilde{x}$ on $[a,b]$ as in this theorem and in Theorem 2.11. Since $\tilde{z} = z + \hat{z}$, it follows that

$$\mathcal{F}_\lambda(\tilde{z}) = \mathcal{F}_\lambda(z) + \mathcal{F}_\lambda(\hat{z}) + 2\int_a^b \Lambda(z,\hat{z})(t)\,\mathrm{d}t, \tag{3.6}$$

where $\Lambda(z,\hat{z})(t)$ is defined in (2.9). Formula (2.11) of Lemma 2.13 yields

$$\mathcal{F}_\lambda(\hat{z}) = \mathcal{F}_0(\hat{z}) - \lambda\,\langle \hat{z}, \hat{z}\rangle_W = \sum_{i=1}^m (\lambda_i - \lambda)\,\beta_i^2, \tag{3.7}$$

while from Lemma 2.12 (with z_i instead of $\hat{z}$) and using $x(a) = 0 = x(b)$ we get

$$\begin{aligned}\int_a^b \Lambda(z,\hat{z})(t)\,\mathrm{d}t &= \sum_{i=1}^m \beta_i \int_a^b \Lambda(z,z_i)(t)\,\mathrm{d}t = \sum_{i=1}^m \beta_i \left(\{x^T u_i\}(t)\,\big|_a^b - \int_a^b \{x^T(u_i' - Cx_i + A^T u_i)\}(t)\,\mathrm{d}t \right) \\ &\overset{(3.1)}{=} \sum_{i=1}^m \beta_i\,\lambda_i\,\langle z, z_i\rangle_W = 0,\end{aligned} \tag{3.8}$$

where the last equality follows from the orthogonality of z and $z_1, \ldots, z_m$. Hence, by inserting formulas (3.5), (3.7), and (3.8) into equation (3.6), we obtain

$$\mathcal{F}_\lambda(z) = \mathcal{F}_\lambda(\tilde{z}) - \mathcal{F}_\lambda(\hat{z}) - 2\int_a^b \Lambda(z,\hat{z})(t)\,\mathrm{d}t \geq \int_a^b \{\tilde{w}^T B \tilde{w}\}(t)\,\mathrm{d}t - \sum_{i=1}^m (\lambda_i - \lambda)\,\beta_i^2,$$

which is what we needed to prove inequality (3.4). □

4. Proof of the Rayleigh principle and applications

Let $\big(\hat{X}(\cdot,\lambda), \hat{U}(\cdot,\lambda)\big)$ be the principal solution of (H_λ), *i.e.*, $\hat{X}(a,\lambda) \equiv 0$ and $\hat{U}(a,\lambda) \equiv I$ for $\lambda \in \mathbb{R}$. Let $n_1(\lambda)$ and $n_2(\lambda)$ denote the number of proper focal points of $\big(\hat{X}(\cdot,\lambda), \hat{U}(\cdot,\lambda)\big)$ in $(a,b]$ and the number of finite eigenvalues of (E) in $(-\infty,\lambda]$, respectively, as in (2.6) and (2.7).

Proof of Theorem 1.1. By Proposition 2.2, condition (1.2) holds, because the functional $\mathcal{F}_{\underline{\lambda}}$ is positive definite for some $\underline{\lambda}$. We start by applying Theorem 2.1, which yields that the matrix $\hat{X}(\cdot,\lambda)$ from the principal solution $\big(\hat{X}(\cdot,\lambda), \hat{U}(\cdot,\lambda)\big)$ of (H_λ) has piecewise constant kernel on $[a,b]$ for every $\lambda \in \mathbb{R}$. Next, since we assume that $\mathcal{F}_{\underline{\lambda}}$ is positive definite for $\underline{\lambda}$ sufficiently negative, then by the oscillation theorem (Thm. 2.9) equality (2.8) holds, and by Proposition 2.2 the principal solution $\big(\hat{X}(\cdot,\underline{\lambda}), \hat{U}(\cdot,\underline{\lambda})\big)$ of ($\mathrm{H}_{\underline{\lambda}}$) has no proper focal points in $(a,b]$.

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Let $m \in \mathbb{N} \cup \{0\}$. Consider the first $m+1$ finite eigenvalues $\lambda_1 \le \cdots \le \lambda_{m+1}$ of (E) with the corresponding orthonormal finite eigenfunctions $z_1, \dots, z_{m+1}$. For convenience we put $\lambda_0 := -\infty$. Suppose first that $\lambda \in (\lambda_m, \lambda_{m+1})$, *i.e.*, $n_2(\lambda) = m$ and λ is not a finite eigenvalue of (E). Then, by (2.4),

$$\operatorname{rank} \hat{X}(b,\lambda) = r(b) = \max_{\mu \in \mathbb{R}} \operatorname{rank} \hat{X}(b,\mu),$$

which in turn implies that

$$n - \operatorname{def} \hat{X}(b,\lambda) = \operatorname{rank} \hat{X}(b,\lambda) = r(b) = \operatorname{rank} \hat{X}(b,\underline{\lambda}) = n - \operatorname{def} \hat{X}(b,\underline{\lambda}).$$

Hence, $\operatorname{def} \hat{X}(b,\lambda) = \operatorname{def} \hat{X}(b,\underline{\lambda})$. This yields that b is not a proper focal point of the principal solution $\big(\hat{X}(\cdot,\lambda), \hat{U}(\cdot,\lambda)\big)$. Consequently, there are exactly $n_1(\lambda)$ proper focal points (including multiplicities) in the open interval (a,b), and $n_1(\lambda) = n_2(\lambda) = m$, by (2.8). Let us denote these proper focal points by $a < \tau_1 < \cdots < \tau_l < b$ with the corresponding multiplicities $m_1, \dots, m_l$, where $\sum_{j=1}^{l} m_j = m$. By the definition of a proper focal point, in particular by (2.3), we have

$$m_j = \dim \big([\operatorname{Ker} \hat{X}(\tau_j^-,\lambda)]^\perp \cap \operatorname{Ker} \hat{X}(\tau_j,\lambda) \big) = \operatorname{def} \hat{X}(\tau_j,\lambda) - \operatorname{def} \hat{X}(\tau_j^-,\lambda), \quad j = 1, \dots, l. \tag{4.1}$$

Consider now a linear combination $\hat{z} = (\hat{x}, \hat{u})$ of the finite eigenfunctions $z_1, \dots, z_m$, that is, $\hat{z} = \sum_{i=1}^{m} \beta_i z_i$ for at this moment unspecified coefficients $\beta_1, \dots, \beta_m \in \mathbb{R}$. Then, by (3.1)–(3.2), $\hat{z}$ is admissible and $\hat{x}(a) = 0 = \hat{x}(b)$.

For the function $\tilde{z} = (\tilde{x}, \tilde{u}) := \hat{z}$ we consider the homogeneous system of linear equations determined by the conditions

$$\tilde{x}(\tau_j) \in \big([\operatorname{Ker} \hat{X}^T(\tau_j^-,\lambda)]^\perp \cap \operatorname{Ker} \hat{X}^T(\tau_j,\lambda) \big)^\perp, \quad j = 1, \dots, l. \tag{4.2}$$

Here $\beta_1, \dots, \beta_m$ are the unknown variables. Since for every $j \in \{1, \dots, l\}$

$$\begin{aligned} \operatorname{def} \hat{X}^T(\tau_j,\lambda) - \operatorname{def} \hat{X}^T(\tau_j^-,\lambda) &= \operatorname{rank} \hat{X}^T(\tau_j^-,\lambda) - \operatorname{rank} \hat{X}^T(\tau_j,\lambda) \\ &= \operatorname{rank} \hat{X}(\tau_j^-,\lambda) - \operatorname{rank} \hat{X}(\tau_j,\lambda) = \operatorname{def} \hat{X}(\tau_j,\lambda) - \operatorname{def} \hat{X}(\tau_j^-,\lambda) = m_j, \end{aligned}$$

and since the points $\tau_1, \dots, \tau_l \in (a,b)$, it follows from (4.1) that there are exactly $\sum_{j=1}^{l} m_j = m$ linear and homogeneous equations in system (4.2) for the m variables $\beta_1, \dots, \beta_m$.

We proceed by showing by induction with respect to $j \in \{0, 1, \dots, l+1\}$ that

$$\tilde{x}(t) \in \operatorname{Im} \hat{X}(t,\lambda) \quad \text{for all } t \in [a,b], \tag{4.3}$$

where we set $\tau_0 := a$ and $\tau_{l+1} := b$. To start the induction, we have $\tilde{x}(\tau_0) = 0 \in \operatorname{Im} \hat{X}(\tau_0,\lambda)$. Suppose now that for some index $j \in \{0, \dots, l\}$ we have $\tilde{x}(t) \in \operatorname{Im} \hat{X}(t,\lambda)$ for all $t \in [a,\tau_j]$. Then, by (2.2), the kernel condition

$$\operatorname{Ker} \hat{X}(t,\lambda) \subseteq \operatorname{Ker} \hat{X}(\tau,\lambda) \quad \text{for all } t, \tau \in [\tau_j, \tau_{j+1}),\ \tau \le t,$$

is satisfied, so that Lemma 2.5 on $[\tau_j, \tau_{j+1})$ (more precisely, on $[\tau_j, s]$ for every $s \in (\tau_j, \tau_{j+1})$) and the induction hypothesis yield the image condition

$$x(t) \in \operatorname{Im} \hat{X}(t,\lambda) \quad \text{for all } t \in [a, \tau_{j+1}). \tag{4.4}$$

If we prove that also $\tilde{x}(\tau_{j+1}) \in \hat{X}(\tau_{j+1},\lambda)$, then the image condition (4.3) will be established. To this end, we first suppose that $j < l$. Then, by (4.2) and the De Morgan law for the orthogonal complement of the intersection of two subspaces, we have

$$\tilde{x}(\tau_{j+1}) \in \operatorname{Ker} \hat{X}^T(\tau_{j+1}^-,\lambda) + [\operatorname{Ker} \hat{X}^T(\tau_{j+1},\lambda)]^\perp. \tag{4.5}$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Since the continuity of $\hat{X}^T(\cdot)$ yields $\operatorname{Ker}\hat{X}^T(\tau_{j+1}^-,\lambda)\subseteq\operatorname{Ker}\hat{X}^T(\tau_{j+1},\lambda)$, *i.e.*,

$$[\operatorname{Ker}\hat{X}^T(\tau_{j+1},\lambda)]^\perp\subseteq[\operatorname{Ker}\hat{X}^T(\tau_{j+1}^-,\lambda)]^\perp,$$

it follows that the sum of the subspaces in (4.5) is a direct sum. Furthermore, by the continuity of $\tilde{x}(\cdot)$ and the already proven image condition (4.4), we also have

$$\tilde{x}(\tau_{j+1})=\tilde{x}(\tau_{j+1}^-)\in\operatorname{Im}\hat{X}(\tau_{j+1}^-,\lambda)=[\operatorname{Ker}\hat{X}^T(\tau_{j+1}^-,\lambda)]^\perp. \tag{4.6}$$

Hence, from (4.5) and (4.6) we obtain

$$\tilde{x}(\tau_{j+1})\in[\operatorname{Ker}\hat{X}^T(\tau_{j+1},\lambda)]^\perp=\operatorname{Im}\hat{X}(\tau_{j+1},\lambda),$$

which is what we wanted to prove. On the other hand, if $j=l$ holds, then by using $\tau_{l+1}=b$ and the definition of $\tilde{x}(b)$ we get $\tilde{x}(\tau_{l+1})=\tilde{x}(b)=0\in\operatorname{Im}\hat{X}(\tau_{l+1},\lambda)$. The image condition (4.3) is therefore established.

We now apply the extended global Picone formula (Thm. 3.1) with $z:=0$ to get

$$0=\mathcal{F}_\lambda(z)\geq\int_a^b\{\tilde{w}^TB\tilde{w}\}\,(t)\,\mathrm{d}t+\sum_{i=1}^m(\lambda-\lambda_i)\,\beta_i^2\geq 0,$$

because (1.2) is assumed and $\lambda>\lambda_i$ for all $i=1,\dots,m$. Consequently, $\beta_1=\cdots=\beta_m=0$, so that the linear system representing conditions (4.2) possesses only the trivial solution. In turn, the coefficient matrix of the system in (4.2) is invertible.

Let now $z=(x,u)$ be admissible with $x(a)=0=x(b)$ and $z\perp z_1,\dots,z_m$. Then for $\tilde{z}:=z+\hat{z}$ the conditions in (4.2) represent a linear system for $\beta_1,\dots,\beta_m$ (in general this system may be nonhomogeneous) with invertible coefficient matrix, as we just proved. Therefore, there exist unique $\beta_1,\dots,\beta_m\in\mathbb{R}$ satisfying the linear system in (4.2), which implies as in the previous part of the proof that the image condition (4.3) holds for this $\tilde{z}=(\tilde{x},\tilde{u})$. Hence, by the extended global Picone formula (Thm. 3.1) and assumption (1.2),

$$\mathcal{F}_\lambda(z)\geq\int_a^b\{\tilde{w}^TB\tilde{w}\}\,(t)\,\mathrm{d}t+\sum_{i=1}^m(\lambda-\lambda_i)\,\beta_i^2\geq 0$$

due to $\lambda>\lambda_i$ for all $i=1,\dots,m$. This yields that

$$\mathcal{F}_0(z)\geq\lambda\,\langle z,z\rangle_W\quad\text{for every }\lambda\in(\lambda_m,\lambda_{m+1}). \tag{4.7}$$

If we now take the limit as $\lambda\to\lambda_{m+1}^-$, we obtain from (4.7) that $\mathcal{F}_0(z)\geq\lambda_{m+1}\,\langle z,z\rangle_W$. Now since z_{m+1} is a solution of $(\mathrm{H}_{\lambda_{m+1}})$, it follows that $\mathcal{F}_{\lambda_{m+1}}(z_{m+1})=0$, that is, $\mathcal{F}_0(z_{m+1})=\lambda_{m+1}\,\langle z_{m+1},z_{m+1}\rangle_W=\lambda_{m+1}$. And since $z_{m+1}=(x_{m+1},u_{m+1})$ is admissible, $x_{m+1}(a)=0=x_{m+1}(b)$, and $z_{m+1}\perp z_1,\dots,z_m$, the minimum in (1.8) is indeed attained at $z=z_{m+1}$.

If $\lambda_{m+1}=\cdots=\lambda_{m+p}$ is a multiple finite eigenvalue (with multiplicity $p\geq 2$), then any admissible $z=(x,u)$ with $x(a)=0=x(b)$ and $z\perp z_1,\dots,z_{m+q}$ (for any $1\leq q\leq p$) satisfies automatically $z\perp z_1,\dots,z_m$. Therefore, by the previous argument we have for such z

$$\mathcal{F}_0(z)\geq\lambda_{m+1}\,\langle z,z\rangle_W=\cdots=\lambda_{m+q}\,\langle z,z\rangle_W,\quad 1\leq q\leq p.$$

The proof of the Rayleigh principle in Theorem 1.1 is now complete. □

Next we make a comment about the existence of finitely or infinitely many finite eigenvalues. With the definition of the set $\mathcal{A}$ of admissible functions with Dirichlet boundary conditions in (1.7), we now define the space

$$\mathcal{W}:=\{\,(Wx)(\cdot),\quad z=(x,u)\in\mathcal{A}\,\}.$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Then all the finite eigenfunctions $z_i = (x_i, u_i)$ belong to $\mathcal{W}$ (or more precisely, the functions $(Wx_i)(\cdot)$ belong to $\mathcal{W}$). Therefore, the number of finite eigenvalues cannot be larger than $\dim \mathcal{W}$. Consequently, our Rayleigh principle in Theorem 1.1 yields the following.

Corollary 4.1. *Assume (1.5) and $\mathcal{F}_{\underline{\lambda}}$ is positive definite for some $\underline{\lambda} < 0$.*

(i) *The eigenvalue problem (E) has infinitely many finite eigenvalues $-\infty < \lambda_1 \le \lambda_2 \dots$ with $\lambda_m \to \infty$ as $m \to \infty$ if and only if $\dim \mathcal{W} = \infty$.*

(ii) *The eigenvalue problem (E) has exactly $p \in \mathbb{N} \cup \{0\}$ finite eigenvalues if and only if $\dim \mathcal{W} = p$. In this case for every $m \in \{0, \dots, p\}$ equality (1.8) holds, in which we put $\lambda_{p+1} := \infty$ when $m = p$.*

The above corollary is a generalization of [7], Corollary 7.7.5 (for the Dirichlet boundary conditions), to the case of abnormal linear Hamiltonian systems.

Remark 4.2. If a set $\{\lambda_1 \le \lambda_2 \le \dots\}$ with $\lambda_m \to \infty$ as $m \to \infty$ of finite eigenvalues of (E) satisfies the Rayleigh principle in Theorem 1.1, then it is complete, that is, there are no further finite eigenvalues of (E). This follows from Theorem 1.1 in the same way as in the case of controllable systems in [7], Remark 7.7.2.

The following supplement – the expansion theorem – is a traditional result connected to the Rayleigh principle. Denote by $\|z\|_W := \sqrt{\langle z, z\rangle_W}$ the (semi-) norm in the space $\mathcal{A}$ of admissible functions with Dirichlet boundary conditions induced by the inner product $\langle \cdot, \cdot \rangle_W$. Note that $\|z\|_W$ depends on the x part of $z = (x, u)$ only. The following result is a generalization of [7], Theorem 7.7.6, to abnormal linear Hamiltonian systems.

Theorem 4.3 (expansion theorem)**.** *Assume (1.5) and $\mathcal{F}_{\underline{\lambda}}$ is positive definite for some $\underline{\lambda} < 0$. Denote by $\mathcal{I}$ the index set which is equal to $\mathbb{N}$ if $\dim \mathcal{W} = \infty$ and which is equal to $\{1, \dots, p\}$ if $\dim \mathcal{W} = p \ge 1$. Let $z = (x, u)$ be admissible with $x(a) = 0 = x(b)$. Then*

$$x = \sum_{i \in \mathcal{I}} c_i\, x_i, \quad \text{i.e.,} \quad \lim_{m \to \infty} \Big\| z - \sum_{i=1}^{m} c_i\, z_i \Big\|_W = 0, \quad \text{where } c_i := \langle z, z_i \rangle_W \text{ for every } i \in \mathcal{I}. \tag{4.8}$$

Proof. Let $z = (x, u)$ be admissible with $x(a) = 0 = x(b)$ and define c_i by (4.8) for $i \in \mathcal{I}$.
Case 1 ($\mathcal{I} = \mathbb{N}$)**.** For every index $m \in \mathbb{N}$ we define $\tilde z = (\tilde x, \tilde u)$ by $\tilde z_m := z - \hat z_m$, where $\hat z_m = (\hat x_m, \hat u_m) := \sum_{i=1}^{m} c_i\, z_i$. Then $\tilde z_m$ is admissible, $\tilde x_m(a) = 0 = \tilde x_m(b)$, and $\tilde z_m \perp z_1, \dots, z_m$, because for each $j \in \{1, \dots, m\}$

$$\langle \tilde z_m, z_j \rangle_W = \langle z, z_j \rangle_W - \sum_{i=1}^{m} c_i \langle z_i, z_j \rangle_W = c_j - c_j = 0.$$

Therefore, the Rayleigh principle (Thm. 1.1) yields

$$\mathcal{F}_0(\tilde z_m) \ge \lambda_{m+1} \langle \tilde z_m, \tilde z_m \rangle_W. \tag{4.9}$$

On the other hand,

$$\mathcal{F}_0(\tilde z_m) = \mathcal{F}_0(z) - 2 \sum_{i=1}^{m} c_i \int_a^b \Lambda(z, z_i)(t)\, \mathrm{d}t + \mathcal{F}_0(\hat z_m), \tag{4.10}$$

where, by Lemmas 2.12 and 2.13,

$$\int_a^b \Lambda(z, z_i)(t)\, \mathrm{d}t = \lambda_i \langle z, z_i \rangle_W = \lambda_i\, c_i \quad \text{and} \quad \mathcal{F}_0(\hat z_m) = \sum_{i=1}^{m} \lambda_i\, c_i^2.$$

Hence, from (4.10) we get

$$\mathcal{F}_0(\tilde z_m) = \mathcal{F}_0(z) - \sum_{i=1}^{m} \lambda_i\, c_i^2 \le \mathcal{F}_0(z) - \lambda_1 \sum_{i=1}^{m} c_i^2,$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

which together with inequality (4.9) implies that

$$\lambda_{m+1}\,\langle \tilde{z}_m, \tilde{z}_m\rangle_W \le \mathcal{F}_0(\tilde{z}_m) \le \mathcal{F}_0(z) - \lambda_1 \sum_{i=1}^{m} c_i^2. \tag{4.11}$$

Take now the index m so large that $\lambda_{m+1} > 0$, which is possible due to $\lambda_m \to \infty$. Then from inequality (4.11) we obtain

$$\|z - \hat{z}_m\|_W^2 = \langle \tilde{z}_m, \tilde{z}_m\rangle_W \le \frac{1}{\lambda_{m+1}} \left(\mathcal{F}_0(z) - \lambda_1 \sum_{i=1}^{m} c_i^2 \right) \to 0 \quad \text{as } m \to \infty.$$

Case 2 ($\mathcal{I} = \{1, \dots, p\}$)**.** In this situation the proof is similar to the proof of the expansion theorem in the discrete time case in [1], Theorem 4.7. One first shows by using Corollary 4.1(ii) that any admissible $z = (x, u)$ with $x(a) = 0 = x(b)$ which is orthogonal to all finite eigenfunctions $z_1, \dots, z_p$ satisfies $\mathcal{F}_0(z) \ge \lambda \langle z, z\rangle_W$ for all $\lambda \in \mathbb{R}$. However, this is possible only if $\langle z, z\rangle_W = 0$, *i.e.*, $(Wx)(\cdot) \equiv 0$. Take now a new admissible $z = (x, u)$ with $x(a) = 0 = x(b)$. Then $\tilde{z} = (\tilde{x}, \tilde{u}) := z - \sum_{i=1}^{p} c_i\, z_i$ is admissible, $\tilde{x}(a) = 0 = \tilde{x}(b)$, and $\tilde{z} \perp z_1, \dots, z_p$. Then, by the previous step in the proof, $(W\tilde{x})(\cdot) \equiv 0$, that is, equation (4.8) holds. The proof of the expansion theorem is now complete. □

Remark 4.4. Suppose $\dim \mathcal{W} = p < \infty$ and $m \in \{0, \dots, p\}$. If an admissible z minimizes the Rayleigh quotient in (1.8), that is, if $\mathcal{F}_0(z) = \lambda_{m+1} \langle z, z\rangle_W$ for such an admissible $z = (x, u)$ with $x(a) = 0 = x(b)$ and $z \perp z_1, \dots, z_m$, then for $\tilde{x}(\cdot) := x(\cdot)$ and $\tilde{u}(\cdot) := \sum_{i=1}^{m} c_i\, u_i(\cdot)$ the pair $\tilde{z} = (\tilde{x}, \tilde{u})$ satisfies the Euler equation

$$\tilde{u}'(t) = C(t)\,\tilde{x}(t) - A^T(t)\,\tilde{u}(t) - \lambda_{m+1} W(t)\,\tilde{x}(t), \quad B(t)\,\tilde{u}(t) = B(t)\,u(t), \quad t \in [a, b].$$

The proof is the same as in the discrete case in [1], Theorem 4.8, and it is hence omitted.

Our second application of the Rayleigh principle concerns the coercivity of the quadratic functional $\mathcal{F}_0$. The functional $\mathcal{F}_0$ is *coercive* if there exists $\alpha > 0$ such that

$$\mathcal{F}_0(z) \ge \alpha \int_a^b |x(t)|^2\, \mathrm{d}t \quad \text{for every admissible } z = (x, u) \text{ with } x(a) = 0 = x(b).$$

As we shall see, it follows from Theorem 1.1 that $\mathcal{F}_0$ is coercive if and only if the eigenvalue problem

$$\left.\begin{aligned} x' = A(t)\,x + B(t)\,u, \quad u' = C(t)\,x - A^T(t)\,u - \lambda\, x, \quad t \in [a, b] \\ x(a) = 0 = x(b), \end{aligned}\right\} \tag{E_1}$$

that is, the eigenvalue problem (E) with $W(\cdot) \equiv I$, has the smallest eigenvalue $\lambda_1 > 0$ (so that we can take $\alpha := \alpha_1$ in the definition of the coercivity of $\mathcal{F}_0$). Note that since $W(\cdot) \equiv I$ is now positive definite, then the assumption on $\mathcal{F}_{\underline{\lambda}} > 0$ for some $\underline{\lambda} < 0$ in Theorem 1.1 can be dropped, as the statement of Lemma 2.10 asserts.

Theorem 4.5 (coercivity)**.** *The functional $\mathcal{F}_0$ is positive definite if and only if it is coercive.*

Proof. We only need to show that $\mathcal{F}_0 > 0$ implies the coercivity of $\mathcal{F}_0$, since the opposite implication holds trivially. Therefore, assume that $\mathcal{F}_0$ is positive definite. Then for any admissible $z = (x, u)$ with $x(a) = 0 = x(b)$ and $x(\cdot) \not\equiv 0$ we have $\mathcal{F}_0(z)/\langle z, z\rangle > 0$, where $\langle z, z\rangle := \langle z, z\rangle_W$ for our matrix $W(\cdot) \equiv I$ in (E_1). Therefore, by the Rayleigh principle (Thm. 1.1) applied to the eigenvalue problem (E_1), the smallest finite eigenvalue λ_1 of (E_1) satisfies $\lambda_1 \ge 0$. If $\lambda_1 = 0$, then there exists a corresponding finite eigenfunction $z_1 = (x_1, u_1)$ with $x_1(a) = 0 = x_1(b)$ and $x_1(\cdot) \not\equiv 0$, which is admissible and for which $\mathcal{F}_0(z_1) = \lambda_1 \langle z_1, z_1\rangle = 0$. This, however, contradicts the assumed positivity of $\mathcal{F}_0$. Therefore, we must have $\lambda_1 > 0$. The Rayleigh principle then yields that $\mathcal{F}_0(z) \ge \lambda_1 \langle z, z\rangle$ for every admissible $z = (x, u)$ with $x(a) = 0 = x(b)$, or in other words, the functional $\mathcal{F}_0$ is coercive with $\alpha := \lambda_1$. □

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

In the proof of [15], Lemma 2.4, p. 231, Wahrheit discusses the notion of a right proper focal point for a conjoined basis (X,U) of (H). Although this notion has not been used anywhere else than in that proof, it seems to be a natural dual version of the (left) proper focal point notion from Section 2.1. More precisely, we say that a point $t_0 \in [a,b)$ is a *right proper focal point* of the conjoined basis (X,U) of (H), provided $\operatorname{Ker} X(t_0^+) \subsetneqq \operatorname{Ker} X(t_0)$, and then the number $m := \operatorname{def} X(t_0) - \operatorname{def} X(t_0^+)$ is its multiplicity. Following (2.3) we then have

$$m = \dim\big([\operatorname{Ker} X(t_0^+)]^\perp \cap \operatorname{Ker} X(t_0)\big). \tag{4.12}$$

The nonexistence of these right proper focal points in $[a,b)$ means that the kernel of $X(\cdot)$ is nondecreasing (measuring in the direction from a to b), *i.e.*,

$$\operatorname{Ker} X(t) \subseteq \operatorname{Ker} X(\tau) \quad \text{for all } t,\tau \in [a,b],\ t \le \tau. \tag{4.13}$$

The kernel condition in (4.13) is shown to be equivalent to the positivity of $\mathcal{F}_0$ in [6], Corollary 6.1, which is a counterpart of Proposition 2.2. If we define the principal solution $(\hat X_a, \hat U_a)$ of (H) at a by the initial conditions $\hat X_a(a) = 0$ and $\hat U_a(a) = I$, that is, $(\hat X_a, \hat U_a) \equiv (\hat X, \hat U)$, and the principal solution $(\hat X_b, \hat U_b)$ of (H) at b by the initial conditions $\hat X_b(b) = 0$ and $\hat U_b(b) = -I$, then we get the following equivalence.

Corollary 4.6. *Assume* (1.2). *The principal solution of* (H) *at a has no (left) proper focal points in $(a,b]$ if and only if the principal solution of* (H) *at b has no (right) proper focal points in $[a,b)$.*

Remark 4.7. From this perspective, *all the theory* of linear Hamiltonian systems without normality, including the oscillation theorems in Appendix B and the Rayleigh principle in Theorem 1.1 can be easily formulated and proven by using this alternative notion of right proper focal points. In addition, the transformation relating the corresponding results in both theories (of left proper focal points and right proper focal points) is $t \mapsto a+b-t$ for $t \in [a,b]$. Note also that the distinction between the left and right proper focal points of (X,U) is possible only in the abnormal case, because under normality the left and right multiplicities of t_0 in (2.3) and (4.12) are both equal to $\operatorname{def} X(t_0)$ (see also Sect. 1).

An application of the corresponding oscillation theorem to this setting then leads, for example, to the following improvement of Corollary 4.6.

Corollary 4.8. *Assume* (1.2) *and let $m \in \mathbb{N} \cup \{0\}$ be fixed. The principal solution of* (H) *at a has m (left) proper focal points in $(a,b]$ if and only if the principal solution of* (H) *at b has m (right) proper focal points in $[a,b)$.*

The statement in Corollary 4.8 is a continuous time version of the discrete result in [4], Theorem 1.1.

Appendix A. Geometric characterization of finite eigenvalues

In this section we shall prove that the algebraic and geometric multiplicities of finite eigenvalues of (E) are the same, *cf.* Remark 2.7 (i). First we need an auxiliary lemma.

Lemma A.1. *Assume* (1.5) *and let $\big(X(\cdot,\lambda), U(\cdot,\lambda)\big)$ be a conjoined basis of* (H_λ) *such that the initial conditions $X(a,\lambda)$ and $U(a,\lambda)$ do not depend on λ,* i.e., *condition* (B.1) *of Appendix B holds, and fix $t \in [a,b]$. Then the kernel of $X(t,\cdot)$ is piecewise constant with respect to λ on $\mathbb{R}$. More precisely, for every $\lambda_0 \in \mathbb{R}$ there exists $\delta > 0$ such that*

$$\operatorname{Ker} X(t,\lambda) = \operatorname{Ker} X(t,\lambda_0^+) = \operatorname{Ker} X(t,\lambda_0^-) \quad \textit{for all } \lambda \in (\lambda_0-\delta, \lambda_0+\delta) \setminus \{\lambda_0\}. \tag{A.1}$$

Proof. Let $\lambda_0 \in \mathbb{R}$ and choose a conjoined basis $\big(\tilde X(\cdot,\lambda), \tilde U(\cdot,\lambda)\big)$ of (H_λ) according to [7], Proposition 4.1.1 and Theorem 3.1.2, such that $(\tilde X, \tilde U)$ and (X,U) are normalized conjoined bases, $\tilde X(t,\lambda_0)$ is invertible, the matrix $(\tilde X^{-1}X)(t,\lambda_0) \ge 0$, and the initial conditions $\tilde X(a,\lambda)$ and $\tilde U(a,\lambda)$ do not depend on λ. Put $X(\lambda) := X(t,\lambda)$ and $\tilde X(\lambda) := \tilde X(t,\lambda)$.

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Then by continuity, assumption (1.5), and [7], Lemma 4.1.4, there exists $\varepsilon > 0$ such that

$$\tilde{X}(\lambda) \text{ is invertible and } (\tilde{X}^{-1}X)(\lambda) \text{ is nondecreasing on } (\lambda_0 - \varepsilon, \lambda_0 + \varepsilon).$$

Let $\mu_1(\lambda) \le \cdots \le \mu_n(\lambda)$ denote the eigenvalues of the symmetric matrix $(\tilde{X}^{-1}X)(\lambda)$, so that

$$0 \le \mu_1(\lambda) \le \cdots \le \mu_n(\lambda) \quad \text{for } \lambda \in [\lambda_0, \lambda_0 + \varepsilon).$$

If $\lambda \in [\lambda_0, \lambda_0 + \varepsilon)$ and $c \in \operatorname{Ker} X(\lambda)$, then $c^T(\tilde{X}^{-1}X)(\lambda)\, c = 0$, and by the monotonicity of $(\tilde{X}^{-1}X)(\cdot)$, we have $0 \le c^T(\tilde{X}^{-1}X)(\lambda_0)\, c \le c^T(\tilde{X}^{-1}X)(\nu)\, c \le c^T(\tilde{X}^{-1}X)(\lambda)\, c = 0$, so that $c \in \operatorname{Ker} X(\nu)$ for all $\nu \in [\lambda_0, \lambda]$. Hence, we proved that $\operatorname{Ker} X(\lambda) \subseteq \operatorname{Ker} X(\nu)$ for $\lambda_0 \le \nu \le \lambda < \lambda_0 + \varepsilon$. Therefore, there exists $\delta \in (0, \varepsilon)$ such that

$$\operatorname{Ker} X(\lambda) \equiv \operatorname{Ker} X(\lambda_0 + \delta) = \operatorname{Ker} X(\lambda_0^+) \quad \text{for all } \lambda \in (\lambda_0, \lambda_0 + \delta),$$

and similarly

$$\operatorname{Ker} X(\lambda) \equiv \operatorname{Ker} X(\lambda_0 - \delta) = \operatorname{Ker} X(\lambda_0^-) \quad \text{for all } \lambda \in (\lambda_0 - \delta, \lambda_0).$$

Since the function $(\tilde{X}^{-1}X)(\lambda)\, c$ is entire in its argument λ for every $c \in \mathbb{R}^n$, it follows that either $(\tilde{X}^{-1}X)(\lambda)\, c \equiv 0$ in $(\lambda_0 - \delta, \lambda_0 + \delta) \setminus \{\lambda_0\}$ or $(\tilde{X}^{-1}X)(\lambda)\, c \ne 0$ for all $\lambda \in (\lambda_0 - \delta, \lambda_0 + \delta) \setminus \{\lambda_0\}$. Hence, equality (A.1) holds. □

Now we state and prove the main result of this section. Here $\big(\hat{X}(\cdot, \lambda), \hat{U}(\cdot, \lambda)\big)$ denotes the principal solution of (H_λ) at a as in Definition 2.6.

Theorem A.2 (geometric characterization of finite eigenvalues)**.** *Assume* (1.5). *The number $\lambda_0 \in \mathbb{R}$ is a finite eigenvalue of the eigenvalue problem* (E) *with (algebraic) multiplicity $\theta(\lambda_0) = r(b) - \operatorname{rank} \hat{X}(b, \lambda_0) \ge 1$ if and only if there exists a corresponding finite eigenfunction $z(\cdot, \lambda_0) = (x(\cdot, \lambda_0), u(\cdot, \lambda_0))$ of* (E) *with $W(\cdot)\, x(\cdot, \lambda_0) \not\equiv 0$ on $[a, b]$. In this case the dimension of the corresponding eigenspace,* i.e.*, the geometric multiplicity of λ_0, equals $\theta(\lambda_0)$. That is, the dimension of the space given in* (2.5) *is $\theta(\lambda_0)$.*

Proof. By Lemma A.1 with $t := b$ and $\big(X(\cdot, \lambda), U(\cdot, \lambda)\big) := \big(\hat{X}(\cdot, \lambda), \hat{U}(\cdot, \lambda)\big)$, the set

$$\mathcal{V} := \operatorname{Ker} \hat{X}(b, \lambda^+) = \operatorname{Ker} \hat{X}(b, \lambda^-) \quad \text{does not depend on } \lambda \in \mathbb{R}.$$

Now $z = (x, u)$ solves (H_{λ_0}) with $x(a) = 0$ if and only if

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \hat{X}(t, \lambda_0)\, c \\ \hat{U}(t, \lambda_0)\, c \end{pmatrix} \quad \text{for all } t \in [a, b] \text{ for some } c \in \mathbb{R}^n.$$

Then $x(b) = 0$ means that $\hat{X}(b, \lambda_0)\, c = 0$, *i.e.*, $c \in \operatorname{Ker} \hat{X}(b, \lambda_0)$.

First suppose that $z = (x, u)$ solves (H_{λ_0}) with $x(a) = 0$ and with

$$W(t)\, x(t) = 0 \quad \text{on } [a, b].$$

Then, for some $c \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$ we have

$$\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} \hat{X}(t, \lambda) \\ \hat{U}(t, \lambda) \end{pmatrix} c = \begin{pmatrix} \hat{X}(t, \lambda_0) \\ \hat{U}(t, \lambda_0) \end{pmatrix} c \quad \text{on } [a, b],$$

because these functions solve the same initial value problem. Hence, $c \in \mathcal{V}$.

Next, let $c \in \mathcal{V}$ be given. Then $\hat{X}(b, \lambda)\, c = 0$ for all $\lambda \in \mathbb{R}$ and

$$x(t) := [\hat{X}(t, \lambda) - \tilde{X}(t, \lambda)\, (\tilde{X}^{-1}X)(b, \lambda)]\, c = \hat{X}(t, \lambda)\, c \quad \text{on } [a, b]$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

satisfies $x(b) = \hat{X}(b,\lambda)\,c = 0$ for all $\lambda \in \mathbb{R}$, where $\big(\tilde{X}(\cdot,\lambda), \tilde{U}(\cdot,\lambda)\big)$ denotes a conjoined basis of (H_λ) such that $\tilde{X}(a,\lambda)$ and $\tilde{U}(a,\lambda)$ do not depend on λ, $(\tilde{X},\tilde{U})$ and $(\hat{X},\hat{U})$ are normalized conjoined bases, and where $\tilde{X}(b,\lambda_0)$ is invertible (similarly as in the proof of Lem. A.1). Then, by [7], Lemma 4.1.4,

$$0 = \frac{\mathrm{d}}{\mathrm{d}\lambda}\, c^T(\tilde{X}^{-1}X)(b,\lambda)\,c = \int_a^b x^T(t)\,W(t)\,x(t)\,\mathrm{d}t,$$

so that $W(t)\,x(t) = 0$ on $[a,b]$ by (1.5).

Thus, we have shown that a nonzero element of the eigenspace (2.5) of the finite eigenvalue λ_0 is of the form $W(\cdot)\,\hat{X}(\cdot,\lambda_0)\,c$ for some $c \in \operatorname{Ker}\hat{X}(b,\lambda_0)\setminus\mathcal{V}$. Then we obtain from Definition 2.6 that the geometric multiplicity of λ_0 is equal to (note that $\dim\mathcal{V} = n - r(b)$)

$$\dim\operatorname{Ker}\hat{X}(b,\lambda_0) - \dim\mathcal{V} = n - \operatorname{rank}\hat{X}(b,\lambda_0) - n + r(b) = r(b) - \operatorname{rank}\hat{X}(b,\lambda_0) = \theta(\lambda_0),$$

which proves the theorem. □

Appendix B. Oscillation theorems

In this section we develop the tools which we need in order to prove Theorem 2.9. The proofs in this section (as well as in the previous section) are extracted from the paper [10] and specified to the continuous time setting. This yields a considerable simplification of these proofs. In addition, the central assumptions in [10], equation (3.7), now reduce to the assumption (1.2). Moreover, the proofs presented here turn out to be simpler and shorter than the proofs of the oscillation theorems in [14,15].

Let be given a conjoined basis $\big(X(\cdot,\lambda), U(\cdot,\lambda)\big)$ of (H_λ) such that the initial conditions $X(a,\lambda)$ and $U(a,\lambda)$ do not depend on λ, *i.e.*,

$$X(a,\lambda) \equiv X(a), \quad U(a,\lambda) \equiv U(a). \tag{B.1}$$

Under (1.2) and counting the multiplicities, we denote by

$$n_1(\lambda) := \text{ the number of proper focal points of } \big(X(\cdot,\lambda), U(\cdot,\lambda)\big) \text{ in } (a,b], \tag{B.2}$$

and as before we define

$$r(t) := \max_{\nu\in\mathbb{R}} \operatorname{rank} X(t,\nu).$$

Then by using that $X(t,\cdot)$ is an entire function in λ for a fixed t, we have for all $\lambda \in \mathbb{R}$

$$r(t) = \operatorname{rank} X(t,\lambda^+) = \operatorname{rank} X(t,\lambda^-), \quad t \in (a,b], \tag{B.3}$$

$$r(a) \equiv \operatorname{rank} X(a,\lambda), \quad \text{under (B.1)}. \tag{B.4}$$

Our first result in this section is the local oscillation theorem.

Theorem B.1 (local oscillation theorem)**.** *Assume* (1.2) *and* (1.5) *and let* $\big(X(\cdot,\lambda), U(\cdot,\lambda)\big)$ *be a conjoined basis of* (H_λ) *satisfying* (B.1)*. Then for all* $\lambda \in \mathbb{R}$ *we have*

$$n_1(\lambda^+) = n_1(\lambda) < \infty,$$
$$n_1(\lambda^+) - n_1(\lambda^-) = r(b) - \operatorname{rank} X(b,\lambda) \geq 0.$$

Hence, the function $n_1(\cdot)$ *is nondecreasing on* $\mathbb{R}$*, the limit*

$$m := \lim_{\lambda\to-\infty} n_1(\lambda)$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

exists with $m \in \mathbb{N} \cup \{0\}$*, so that for a suitable* $\lambda_0 \in \mathbb{R}$*,* $\lambda_0 < 0$*, we have*

$$n_1(\lambda) \equiv m \quad \text{and} \quad r(b) - \operatorname{rank} X(b,\lambda) \equiv 0 \quad \text{for all } \lambda \le \lambda_0.$$

For the proof of this theorem we need some auxiliary lemmas. In the first lemma, let $\lambda \in \mathbb{R}$ be fixed, and therefore can omit the variable λ in the Hamiltonian system and its solutions, because for the given value of λ the system (H_λ) can be identified with a system of the form (H) for a suitable matrix $C(t)$.

Lemma B.2. *Assume* (1.2)*, let* $a \le \alpha < \beta \le b$*, and suppose that* $(\tilde X, \tilde U)$ *is a conjoined basis of* (H) *such that* $(\tilde X, \tilde U)$ *and* (X,U) *are normalized conjoined bases of* (H) *and such that*

$$\tilde X(t) \text{ is invertible for all } t \in [\alpha,\beta].$$

Let $\tilde m$ *denote the number of proper focal points of* (X,U) *in* $(\alpha,\beta]$*. Then* $0 \le \tilde m \le n$*, and*

$$\tilde m = \operatorname{ind}\big(\tilde X^{-1}X\big)(\alpha) - \operatorname{ind}\big(\tilde X^{-1}X\big)(\beta). \tag{B.5}$$

Proof. First we assume that

$$\operatorname{Ker} X(t) \text{ is constant on the open interval } (\alpha,\beta). \tag{B.6}$$

Then, by the definition (*cf.* (2.3)), there is no proper focal point of (X,U) in (α,β), and

$$\tilde m = \dim \operatorname{Ker} X(\beta) - \dim \operatorname{Ker} X(\beta^-).$$

Let $\mu_1(t) \le \cdots \le \mu_n(t)$ denote the eigenvalues of the symmetric matrix $\big(\tilde X^{-1}X\big)(t)$. Then the temporary assumption (B.6) implies

$$\mu_1(t) \le \cdots \le \mu_r(t) < 0 = \mu_{r+1}(t) = \cdots = \mu_{r+s}(t) < \mu_{r+s+1}(t) \le \cdots \le \mu_n(t)$$

for all $t \in (\alpha,\beta)$, where

$$r := \operatorname{ind}\big(\tilde X^{-1}X\big)(\beta^-) = \operatorname{ind}\big(\tilde X^{-1}X\big)(\alpha^+) \quad \text{and} \quad s := \operatorname{def} X(\beta^-) = \operatorname{def} X(\alpha^+),$$

because no eigenvalue $\mu_i(\cdot)$ can change its sign in (α,β) by the intermediate value theorem, since the eigenvalues $\mu_i(\cdot)$ are continuous and $\operatorname{Ker} X(\cdot)$ is constant on (α,β). By continuity (*cf.* Thm. 2.1) and (B.6), we have that

$$\operatorname{Ker} X(\alpha^+) = \operatorname{Ker} X(\beta^-) \subseteq \operatorname{Ker} X(\alpha) \cap \operatorname{Ker} X(\beta),$$

so that $s \le \min\{\operatorname{def} X(\alpha), \operatorname{def} X(\beta)\}$. Next, we have by an easy calculation (*cf.* [7], Cor. 1.1.4) that

$$\big(\tilde X^{-1}X\big)'(t) = \big(\tilde X^{-1}B(\tilde X^T)^{-1}\big)(t) \ge 0 \quad \text{for all } t \in [\alpha,\beta],$$

where we use the crucial assumption (1.2). Hence, the eigenvalues $\mu_i(\cdot)$ are nondecreasing on $[\alpha,\beta]$, which implies that

$$\operatorname{ind}\big(\tilde X^{-1}X\big)(\alpha) = r \quad \text{and} \quad \operatorname{ind}\big(\tilde X^{-1}X\big)(\beta) = r - \tilde m,$$

and this yields our assertion (B.5) under the temporary assumption (B.6).

Next, by (1.2) and Theorem 2.1, we have that $\operatorname{Ker} X(t)$ is piecewise constant on $[\alpha,\beta]$, so that there exists a partition $\alpha = \tau_0 < \tau_1 < \cdots < \tau_{k+1} = \beta$ of $[\alpha,\beta]$ such that

$$\operatorname{Ker} X(t) \text{ is constant on } (\tau_j, \tau_{j+1}) \text{ for every } j \in \{0,\dots,k\}.$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

If $\tilde{m}(j)$ denotes the number of proper focal points of (X,U) in $(\tau_j,\tau_{j+1}]$, then we obtain from the first part *via* telescope summation

$$0 \le \tilde{m} = \sum_{j=0}^{k} \tilde{m}(j) = \sum_{j=0}^{k} \big\{ \operatorname{ind}(\tilde{X}^{-1}X)(\tau_j) - \operatorname{ind}(\tilde{X}^{-1}X)(\tau_{j+1}) \big\} = \operatorname{ind}(\tilde{X}^{-1}X)(\alpha) - \operatorname{ind}(\tilde{X}^{-1}X)(\beta) \le n,$$

which is the statement of this lemma. □

Remark B.3. Under the assumptions of Lemma B.2, the function $\operatorname{ind}(\tilde{X}^{-1}X)(\cdot)$ is nonincreasing on $[\alpha,\beta]$. The above proof shows that $t_0 \in (\alpha,\beta]$ is a proper focal point of the conjoined basis (X,U) if and only if $\operatorname{ind}(\tilde{X}^{-1}X)(t_0^-) > \operatorname{ind}(\tilde{X}^{-1}X)(t_0)$, and in this case the difference $\operatorname{ind}(\tilde{X}^{-1}X)(t_0^-) - \operatorname{ind}(\tilde{X}^{-1}X)(t_0)$ is its multiplicity. Note also that in this approach the assertion $0 \le \tilde{m} \le n$ is trivial and it corresponds to the Sturmian separation theorem [11], p. 366.

Now we vary λ using the notation above.

Lemma B.4. *Assume* (1.2) *and* (1.5), *let* $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big)$ *be a conjoined basis of* (H_λ) *satisfying* (B.1), *and let* $a \le \alpha < \beta \le b$. *Then for all* $\lambda \in \mathbb{R}$ *we have*

$$\begin{aligned} m(\lambda^+) &= m(\lambda) < \infty, \\ m(\lambda^+) - m(\lambda^-) &= r(\beta) - \operatorname{rank} X(\beta,\lambda) - r(\alpha) + \operatorname{rank} X(\alpha,\lambda), \end{aligned}$$

where $m(\lambda)$ *denotes the number of proper focal points of* $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big)$ *in* $(\alpha,\beta]$.

Proof. First we fix $\lambda_0 \in \mathbb{R}$. By continuity and compactness (using also [7], Prop. 4.1.1), there exists $\varepsilon > 0$, a finite partition $\alpha = \tau_0 < \tau_1 < \cdots < \tau_{k+1} = \beta$, and conjoined bases $\big(\tilde{X}_j(\cdot,\lambda),\tilde{U}_j(\cdot,\lambda)\big)$ of (H_λ) such that for every $j \in \{0,\dots,k\}$:

(i) $\tilde{X}_j(a,\lambda) \equiv \tilde{X}_j(a,\lambda_0)$ and $\tilde{U}_j(a,\lambda) \equiv \tilde{U}_j(a,\lambda_0)$ do not depend on λ,
(ii) $\big(\tilde{X}_j(\cdot,\lambda),\tilde{U}_j(\cdot,\lambda)\big)$ and $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big)$ are normalized conjoined bases of (H_λ),
(iii) $\tilde{X}_j(t,\lambda)$ is invertible for all $(t,\lambda) \in [\tau_j,\tau_{j+1}] \times [\lambda_0-\varepsilon,\lambda_0+\varepsilon]$.

Let $m(j,\lambda)$ denote the number of proper focal points of $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big)$ in $(\tau_j,\tau_{j+1}]$, so that

$$m(\lambda) = \sum_{j=0}^{k} m(j,\lambda) \quad \text{for } \lambda \in \mathbb{R} \tag{B.7}$$

is the number of proper focal points of $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big)$ in $(\alpha,\beta]$. Then, by the previous Lemma B.2 and assumption (1.2),

$$0 \le m(j,\lambda) = \operatorname{ind}(\tilde{X}_j^{-1}X)(\tau_j,\lambda) - \operatorname{ind}(\tilde{X}_j^{-1}X)(\tau_{j+1},\lambda) \le n \tag{B.8}$$

for every $j = 0,\dots,k$ and $\lambda \in [\lambda_0-\varepsilon,\lambda_0+\varepsilon]$. Hence, $m(\lambda_0) < \infty$ is finite. Next, fix any $j = 0,\dots,k$ and $t_0 \in [\tau_j,\tau_{j+1}]$, and put $Q(\lambda) := (\tilde{X}_j^{-1}X)(t_0,\lambda)$ for $\lambda \in [\lambda_0-\varepsilon,\lambda_0+\varepsilon]$. By [7], Lemma 4.1.4, and assumption (1.5) (similarly as in the proof of Thm. A.2), the symmetric function $Q(\cdot)$ is nondecreasing and continuous on the interval $[\lambda_0-\varepsilon,\lambda_0+\varepsilon]$. Therefore, the eigenvalues $\mu_1(\lambda) \le \cdots \le \mu_n(\lambda)$ of $Q(\lambda)$ are nondecreasing and continuous on $[\lambda_0-\varepsilon,\lambda_0+\varepsilon]$, too. This implies immediately that

$$\operatorname{ind} Q(\lambda_0) = \operatorname{ind} Q(\lambda_0^+) \quad \text{and} \quad \operatorname{ind} Q(\lambda_0^-) = \operatorname{ind} Q(\lambda_0) + r(t_0) - \operatorname{rank} X(t_0,\lambda_0). \tag{B.9}$$

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

Note that the $\operatorname{Ker} Q(\cdot)$ is piecewise constant on $[\lambda_0-\varepsilon,\lambda_0+\varepsilon]$, by Lemma A.1. We conclude from (B.8) and (B.9) with $t_0:=\tau_j$ and $t_0:=\tau_{j+1}$ that for every $j=0,\dots,k$

$$m(j,\lambda_0^+)=m(j,\lambda_0), \tag{B.10}$$

$$m(j,\lambda_0^+)-m(j,\lambda_0^-)=r(\tau_{j+1})-\operatorname{rank}X(\tau_{j+1},\lambda_0)-r(\tau_j)+\operatorname{rank}X(\tau_j,\lambda_0). \tag{B.11}$$

It follows from equations (B.7) and (B.10) that $m(\lambda_0^+)=m(\lambda_0)$, and from equations (B.7) and (B.11) *via* telescope summation that

$$\begin{aligned} m(\lambda_0^+)-m(\lambda_0^-)&=\sum_{j=0}^{k}\big\{r(\tau_{j+1})-\operatorname{rank}X(\tau_{j+1},\lambda_0)-r(\tau_j)+\operatorname{rank}X(\tau_j,\lambda_0)\big\}\\ &=r(\beta)-\operatorname{rank}X(\beta,\lambda_0)-r(\alpha)+\operatorname{rank}X(\alpha,\lambda_0). \end{aligned}$$

The proof of this lemma is complete. □

Now we are ready to prove the local oscillation theorem.

Proof of Theorem B.1. If $\alpha=a$, then $r(a)=\operatorname{rank}X(a,\lambda)$ for all $\lambda\in\mathbb{R}$ by assumption (B.1) and conclusion (B.4). Hence, Theorem B.1 follows directly from Lemma B.4 with $\alpha:=a$ and $\beta:=b$, because the additional assertions are clear. □

Let now $\big(\hat X(\cdot,\lambda),\hat U(\cdot,\lambda)\big)$ be the principal solution of (H_λ), whose initial conditions do not depend on λ and therefore it satisfies condition (B.1). Let $n_1(\lambda)$ be defined by (B.2) through $\big(\hat X(\cdot,\lambda),\hat U(\cdot,\lambda)\big)$, or in this case $n_1(\lambda)$ is defined by (2.6). Furthermore, let $n_2(\lambda)$ be defined by (2.7).

Theorem B.5 (global oscillation theorem)**.** *Assume* (1.2) *and* (1.5). *Then for all* $\lambda\in\mathbb{R}$

$$\begin{aligned} n_2(\lambda^+)&=n_2(\lambda)<\infty,\\ n_2(\lambda^+)-n_2(\lambda^-)&=n_1(\lambda^+)-n_1(\lambda^-)\geq 0, \end{aligned}$$

and there exists $m\in\mathbb{N}\cup\{0\}$ *such that*

$$n_1(\lambda)=n_2(\lambda)+m\quad \textit{for all } \lambda\in\mathbb{R}. \tag{B.12}$$

Moreover, for a suitable $\lambda_0\in\mathbb{R}$, $\lambda_0<0$, *we have*

$$n_2(\lambda)\equiv 0\quad \textit{and}\quad n_1(\lambda)\equiv m\quad \textit{for all } \lambda\leq\lambda_0. \tag{B.13}$$

Proof. This global oscillation theorem follows immediately from the local oscillation theorem (Thm. B.1), in which we take $\big(X(\cdot,\lambda),U(\cdot,\lambda)\big):=\big(\hat X(\cdot,\lambda),\hat U(\cdot,\lambda)\big)$ to be the principal solution of system (H_λ). □

Proof of Theorem 2.9. From Proposition 2.2 we have that the functional $\mathcal{F}_{\underline{\lambda}}>0$ if and only if the principal solution $\big(\hat X(\cdot,\underline{\lambda}),\hat U(\cdot,\underline{\lambda})\big)$ of ($\mathrm{H}_{\underline{\lambda}}$) has no proper focal points in $(a,b]$. By formula (B.12) of Theorem B.5, this is equivalent to equality (2.8), because in this case we have $m=0$. Equation (B.13) then yields $n_2(\lambda)\equiv 0$ for $\lambda\leq\lambda_0$, so that there are no finite eigenvalues less than λ_0. Hence, the finite eigenvalues of (E) are bounded from below. □

Acknowledgements. The authors wish to thank Masaryk University and the University of Ulm for the hospitality provided while conducting this research project.

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press

References

[1] M. Bohner, O. Došlý and W. Kratz, Sturmian and spectral theory for discrete symplectic systems. *Trans. Am. Math. Soc.* **361** (2009) 3109–3123.

[2] W.A. Coppel, *Disconjugacy, Lecture Notes in Mathematics* **220**. Springer-Verlag, Berlin, Heidelberg (1971).

[3] O. Došlý and W. Kratz, Oscillation theorems for symplectic difference systems. *J. Difference Equ. Appl.* **13** (2007) 585–605.

[4] J.V. Elyseeva, The comparative index and the number of focal points for conjoined bases of symplectic difference systems in Discrete Dynamics and Difference Equations, in *Proceedings of the Twelfth International Conference on Difference Equations and Applications, Lisbon, 2007*, edited by S. Elaydi, H. Oliveira, J.M. Ferreira and J.F. Alves. World Scientific Publishing Co., London (2010) 231–238.

[5] R. Hilscher and V. Zeidan, Riccati equations for abnormal time scale quadratic functionals. *J. Differ. Equ.* **244** (2008) 1410–1447.

[6] R. Hilscher and V. Zeidan, Nabla time scale symplectic systems. *Differ. Equ. Dyn. Syst.* **18** (2010) 163–198.

[7] W. Kratz, *Quadratic Functionals in Variational Analysis and Control Theory*. Akademie Verlag, Berlin (1995).

[8] W. Kratz, An oscillation theorem for self-adjoint differential systems and the Rayleigh principle for quadratic functionals. *J. London Math. Soc.* **51** (1995) 401–416.

[9] W. Kratz, Definiteness of quadratic functionals. *Analysis (Munich)* **23** (2003) 163–183.

[10] W. Kratz, R. Šimon Hilscher and V. Zeidan, Eigenvalue and oscillation theorems for time scale symplectic systems. *Int. J. Dyn. Syst. Differ. Equ.* **3** (2011) 84–131.

[11] W.T. Reid, *Ordinary Differential Equations*. Wiley, New York (1971).

[12] W.T. Reid, *Sturmian Theory for Ordinary Differential Equations*. Springer-Verlag, New York-Berlin-Heidelberg (1980).

[13] R. Šimon Hilscher and V. Zeidan, Picone type identities and definiteness of quadratic functionals on time scales. *Appl. Math. Comput.* **215** (2009) 2425–2437.

[14] M. Wahrheit, *Eigenwertprobleme und Oszillation linearer Hamiltonischer Systeme* [*Eigenvalue Problems and Oscillation of Linear Hamiltonian Systems*]. Ph.D. thesis, University of Ulm, Germany (2006).

[15] M. Wahrheit, Eigenvalue problems and oscillation of linear Hamiltonian systems. *Int. J. Difference Equ.* **2** (2007) 221–244.

https://doi.org/10.1051/cocv/2011104 Published online by Cambridge University Press