Worst Exponential Decay Rate for Degenerate Gradient flows subject to persistent excitation*

Yacine Chitour[‡,♯], Dario Prandi[†,‡] , and Paolo Mason[†,‡]

[†]CNRS

[‡]Laboratoire des Signaux et Systèmes, CentraleSupélec, Gif-sur-Yvette, France

[♯] Université Paris-Sud, Orsay, France

In this paper, we partially solve an open problem posed by A. Rantzer in 1999, cf. [13], which asks to estimate the worst L_2-gain of the time-varying linear control systems $\dot{x}(t) = -c(t)c(t)^T x(t) + u$, where the signal $c(\cdot)$ is subject to a *persistent excitation* (PE) condition stating that there exists positive constants a, b, T such that, for every $t \geq 0$ one has $a\ Id_n \leq \int_t^{t+T} c(s)c(s)^T ds \leq b\ Id_n$. This question is related to estimate the worst rate of exponential decay of $\dot{x}(t) = -c(t)c(t)^T x(t)$ which is a degenerate gradient flow issued from adaptative control theory [3]. We prove that the worst rate of exponential decay is at most of the magnitude $\frac{a}{(1+b^2)T}$, to be compared with lower bounds obtained previously $\frac{a}{(1+nb^2)T}$. We also provide results for more general classes of (PE) signals.

The solution consists in relating the above mentioned problem to optimal control questions and we study in details their optimal solutions exploiting the Pontryagin Maximum Principle.

*This research was partially supported by the iCODE Institute, research project of the IDEX Paris-Saclay, and by the Hadamard Mathematics LabEx (LMH) through the grant number ANR-11-LABX-0056-LMH in the "Programme des Investissements d'Avenir".

Contents

1. Introduction

In this paper, we consider systems of the type

$$\dot{x}(t) = -c(t)c(t)^\top x(t), \tag{DGF}$$

where $x \in \mathbb{R}^n$ and the signal $c : [0, \infty) \to \mathbb{R}^n$ is square integrable and verifies the *persistent excitation* condition. That is, there exist positive constants a, b, T such that,

$$\forall t \geq 0,\ a\,\mathrm{Id}_n \leq \int_t^{T+t} c(s)c(s)^\top ds \leq b\,\mathrm{Id}_n\,. \tag{PE}$$

Here, Id_n denotes the $n \times n$ identity matrix and the inequalities are to be understood in the sense of quadratic forms. Henceforth, we will denote by $\mathcal{C}_n(a, b, T)$ the set of signals satisfying (PE).

The above dynamics appear in the context of adaptive control and identification of (unknown) parameters, and are usually referred to as *degenerate gradient flow systems* (DGF), since the Euclidean norm is decreasing along its trajectories (cf. [2, 4, 8, 14] and references therein). As an immediate consequence, these trajectories are defined on $[0, +\infty)$, and it is well-known that the (PE) condition is equivalent to global exponential stability of (DGF), see, e.g., [2].

The rate of exponential decay of (DGF) for a given signal $c \in \mathcal{C}_n(a, b, T)$ is defined as

$$R(c) := -\limsup_{t\to\infty} \frac{\ln \|\Phi_c(t, 0)\|}{t}, \tag{1}$$

where $\Phi_c(t+s,t)$ denotes the flow (or fundamental matrix) of (DGF) from t to $t+s$. The main object of interest in this paper is the worst-rate of exponential decay associated with $\mathcal{C}_n(a,b,T)$ and defined as

$$R(a,b,T,n) := \inf\{R(c) \mid c \in \mathcal{C}_n(a,b,T)\}, \tag{2}$$

and, in particular, its estimation in terms of the parameters a,b,T,n. We recall that it is well known (cf., e.g., [2, 4, 8, 15]) that

$$R(a,b,T,n) \geq \frac{Ca}{(1+nb^2)T}. \tag{3}$$

for some universal constant $C>0$ (i.e., independent of the parameters a,b,n,T). Our main result is the following, which shows that, for fixed n, (3) is optimal.

Theorem 1. *There exists $C_0>0$ such that, for every $0<a\leq b$, $T>0$ and integer $n\geq 2$, the worst-rate of exponential decay $R(a,b,T,n)$ defined in (2) satisfies*

$$R(a,b,T,n) \leq \frac{C_0 a}{(1+b^2)T}. \tag{4}$$

Remark 2. Theorem 1 shows in particular that $R(a,b,T,n)$ tends to zero as b tends to infinity. This is in accordance with [7], where it is proved that in general there is no convergence to the origin for trajectories of (DGF) if only the left inequality of (PE) holds true, i.e., $b=+\infty$. More precisely, the authors put forward a “freezing” phenomenon by showing that in this case there exist trajectories of (DGF) which converge, as t tends to infinity, to points different from the origin.

1.1. L^2 gain

As a consequence of Theorem 1 and of the arguments to derive it, we provide a partial answer to the first part of Problem 36 given by A. Rantzer, whose statement is the following. For $c \in \mathcal{C}_n(a,b,T)$, consider the control system defined by

$$\dot{x}(t) = -c(t)c(t)^T x(t) + u(t). \tag{5}$$

For $u \in L_2([0,\infty),\mathbb{R}^n)$, let $x_u \in L_2([0,\infty),\mathbb{R}^n)$ be the trajectory of (5) associated with u and starting at the origin. Since trajectories of the uncontrolled dynamics tend to zero exponentially, the input/output map $u \mapsto x_u$ is well-defined as a linear operator of $L_2(0,\infty)$ and its L_2-gain $\gamma(c)$ is finite. Rantzer's question consists in determining

$$\gamma(a,b,T,n) := \sup_{c\in\mathcal{C}_n(a,b,T,n)} \gamma(c).$$

In that direction, we have the following result.

Theorem 3. *There exists $c_1,c_2>0$ such that, for every $0<a\leq b$ $T>0$ and integer $n\geq 2$,*

$$c_1\frac{T(1+b^2)}{a} \leq \gamma(a,b,n,T) \leq c_2\frac{T(1+nb^2)}{a} \tag{6}$$

1.2. Generalized persistent excitation

Recently, there has been an increasing interest in considering more general types of persistent excitation conditions, cf. [6, 12, 10]. We focus on the following *generalised persistent excitation* condition:

$$a_\ell Id_n \le \int_{\tau_\ell}^{\tau_{\ell+1}} c(t)c(t)^\top dt \le b_\ell \, \mathrm{Id}_n, \tag{PEG}$$

where $(a_\ell)_{\ell\in\mathbb{N}}$, $(b_\ell)_{\ell\in\mathbb{N}}$ are sequences of positive numbers, and $(\tau_\ell)_{\ell\in\mathbb{N}}$ is a strictly increasing sequence of positive times such that $\tau_\ell \to +\infty$ as $\ell \to +\infty$.

An important question is then to determine under which condition (PEG) guarantees global asymptotic stability (GAS) for (DGF). In [12] (cf. also [6]) the author proved the sufficient condition:

$$\sum_{\ell=0}^{\infty} \frac{a_\ell}{1+b_\ell^2} = +\infty. \tag{7}$$

As a byproduct of our analysis, we show that this condition is indeed necessary.

Theorem 4. *All systems* (DGF) *that satisfy condition* (PEG) *are GAS if and only if* (7) *holds.*

We stress that our interest lies in the study of systems satisfying (PEG) as a class. That is, the above theorem states that if (7) is not satisfied, then there exists an input signal satisfying (PEG) that is not GAS. However, for a fixed signal satisfying (PEG), condition (7) is not necessary for GAS, as shown in [6, Prop. 7].

1.3. Strategy of proof

We now turn to a brief description of the strategy of proof. The main idea is to consider optimal control problems whose minimal values provide bounds for the worst-rate of exponential decay.

More precisely, since the dynamics in (DGF) are linear in $x \in \mathbb{R}^n$, the system is amenable to be decomposed in spherical coordinates. Thus, letting $x = r\omega$, for $r = \|x\| \in \mathbb{R}_+$ and $\omega = x/\|x\| \in \mathbb{S}^{n-1}$, (DGF) reads as

$$\dot r = -\left(c^\top \omega\right)^2 r, \tag{8}$$

$$\dot \omega = -c^\top \omega\left(c - \left(c^\top\omega\right)\omega\right). \tag{9}$$

For $c \in \mathcal{C}_n(a,b,T)$, consider the control system defined by (5) and let $\phi_c(\cdot,\cdot)$ be the fundamental matrix associated with c, i.e., for every $0 \le s \le t$, $\phi_c(t,s)$ is the value at time t of the solution of $\dot M = -cc^T M$ with initial condition $M(s) = Id_n$. Observe that, for every $x \in \mathbb{R}^n \setminus \{0\}$ and $c \in \mathcal{C}_n(a,b,T)$, if we let $\Phi_c(t,0)x = (r(t),\omega(t))$, then it holds

$$\ln\left(\frac{\|\Phi_c(T+t,0)x\|}{\|\Phi_c(t,0)x\|}\right) = \ln\left(\frac{r(t+T)}{r(t)}\right) = -\int_t^{t+T} \left(c^\top\omega\right)^2 ds, \quad \forall t \ge 0. \tag{10}$$

Since the last term in the above above equation does not depend of r, this suggests to consider the following optimal control problem

$$\min_{\substack{c\in\mathcal{C}_n^0(a,b,T),\\ \omega_0\in\mathbb{S}^{n-1}}} J(c,\omega_0), \quad J(c,\omega_0) := \int_0^T \left(c^\top \omega\right)^2 dt. \tag{OCP}$$

Here, $\mathcal{C}_n^0(a,b,T)$ denotes the set of restrictions of elements of $\mathcal{C}_n(a,b,T)$ to $[0,T]$ and $\omega : [0,T] \to \mathbb{S}^{n-1}$ is an arbitrary trajectory of (9) with initial condition $\omega(0) = \omega_0$.

Let $\mu(a,b,n)$ be the minimal value obtained for (OCP), which we show in Proposition 7 below to be independent of T. In Section 3 we reduce the proof of the main results to the following.

Proposition 5. *There exists a universal constant $C_0 > 0$ such that, for every $0 < a \leq b$ $T > 0$ and integer $n \geq 2$,*

$$\mu(a,b,n) \leq \frac{C_0 a}{1+b^2}. \tag{11}$$

Moreover, there exists a $2T$-periodic control $c_* \in \mathcal{C}_n(2a,2b,2T)$ and initial condition $\omega_0 \in \mathbb{S}^{n-1}$ such that

$$\omega_*(t) = \frac{\Phi_{c_*}(t,0)\omega_0}{\|\Phi_{c_*}(t,0)\omega_0\|} \tag{12}$$

is a $2T$-periodic trajectory and both $t \mapsto c_*|_{[0,T]}(t)$ and $t \mapsto c_*|_{[T,2T]}(t-T)$, together with the respective initial conditions ω_0 and $\omega_*(T)$, are minimizers for (OCP).

The rest of the paper is then devoted to prove the above proposition. We first observe that, due to the monotonicity with respect to the dimension n of the minimal value $\mu(a,b,n)$, for the first part of the statement it is enough to bound $\mu(a,b,2)$. We then apply Pontryagin Maximum Principle, and we explicitly integrate the resulting Hamiltonian system in the two dimensional case, thus obtaining the result. Finally, the proof of the second part of the statement by a detailed analysis in arbitrary dimension $n \geq 2$ of the extremal trajectories associated with (OCP).

1.4. Notations

We use $\lfloor x \rfloor$ to denote the integer part of the real number x. We let Sym_n be the set of $n \times n$ symmetric real matrices, and by Sym_n^+ the subset of non negative ones. Moreover, for $a \leq b$, we use $\mathrm{Sym}_n(a,b)$ to denote the set of matrices $S \in \mathrm{Sym}_n$ such that $a\,\mathrm{Id}_n \leq S \leq b\,\mathrm{Id}_n$ in the sense of quadratic forms. For every positive integer k, we denote by $\mathbb{S}^k$ the unit sphere of $\mathbb{R}^{k+1}$.

Observe that the class $\mathcal{C}_n(a,b,T)$ of signals that satisfy (PE) can be characterized as the class of signals c such that $Q(t) := \int_t^{t+T} c(s)c(s)^\top ds \in \mathrm{Sym}_n(a,b)$ for every $t \geq 0$. This can be relaxed to a PE condition on measurable functions $S : \mathbb{R}_+ \to \mathrm{Sym}_n^+$, by requiring the existence of positive constants a, b, T such that

$$\int_t^{t+T} S(s)\,ds \;\in\; \mathrm{Sym}_n(a,b),\ \forall t \geq 0. \tag{13}$$

We let $\mathrm{Sym}_n^+(a,b,T)$ be the set of functions that satisfy (13). Clearly, $\mathcal{C}_n(a,b,T)$ can be identified with the projectors in $\mathrm{Sym}_n^+(a,b,T)$. Observe that if $S \in \mathrm{Sym}_n^+(a,b,T)$ (resp. $c \in \mathcal{C}_n(a,b,T)$), then the same is true for $USU^\top$ (resp. Uc), for every orthogonal matrix $U \in \mathrm{O}(n)$.

2. Relaxation of the optimal control problem (OCP) and first consequences

We start this section by showing a simple upper bound for the minimal value $\mu(a,b,T,n)$ of (OCP).

Proposition 6. *With the notations above, it holds*

$$\mu(a,b,T,n) \leq a. \tag{14}$$

Proof. It suffices to consider the control $c \in \mathcal{C}_n^0(a,b,T)$ defined by

$$c(t) = \sqrt{an}e_j \quad \text{if } t \in \left[\frac{(j-1)T}{n}, \frac{jT}{n}\right), \qquad j = 1,\ldots,n. \tag{15}$$

Here, $\{e_1,\ldots,e_n\} \subset \mathbb{R}^n$ denotes the canonical basis of $\mathbb{R}^n$. Indeed, for $\omega_0 = e_1$, we have that $\omega \equiv \omega_0$, $J(cc^\top) = a$, and

$$\int_0^T cc^\top \, dt = a\,\mathrm{Id}_n\,. \qquad \square$$

Since we want to apply techniques of optimal control to get a hold on the minimal value $\mu(a,b,T,n)$ of (OCP) (to be proved to be independant of T), we need to establish existence of minimizers for the control problem at hand. However, $\mathcal{C}_n^0(a,b,T)$ is not closed under the weak-$*$ topology, preventing one to directly deduce such an existence from standard theorems. One must therefore relax the problem.

Since the convex hull of $\mathcal{C}_n^0(a,b,T)$ (identified with a set of at most rank-one matrix valued functions) is $\mathrm{Sym}_n^{+,0}(a,b,T)$, this yields the following optimal control problem:

$$\min_{\substack{S\in \mathrm{Sym}_n^{+,0}(a,b,T)\\ \omega_0\in\mathbb{S}^{n-1}}} J(S,\omega_0), \quad J(S,\omega_0) := \int_0^T \omega^\top S\omega\, dt. \tag{Conv}$$

Here, $\mathrm{Sym}_n^{+,0}(a,b,T)$ denotes the set of restrictions of elements of $\mathrm{Sym}_n^+(a,b,T)$ to $[0,T]$, and ω satisfy

$$\dot\omega = -S\omega + (\omega^\top S\omega)\omega, \qquad \omega(0) = \omega_0. \tag{16}$$

Clearly, $\mu(a,b,T,n)$ coincides with the minimal value obtained for (Conv). We then have the following.

Proposition 7. *The optimal control problem* (Conv) *admits minimizers with constant trace. Moreover, the minimal value $\mu(a,b,T,n)$ is independent of $T > 0$.*

Remark 8. Due to the above, we will henceforth let $\mu(a,b,n) = \mu(a,b,T,n)$.

Proof. We need to compactify (Conv). For that purpose, we consider the following optimal control problem

$$\begin{cases} \min\{L(S,\omega_0) \mid \mathcal{T} \in [an,bn],\, S \in \mathrm{Sym}^{+,0}_{n,1}(a,b,\mathcal{T}), \quad \omega_0 \in \mathbb{S}^{n-1}\}, \\ L(S,\omega_0) := \displaystyle\int_0^{\mathcal{T}} \omega(t)^\top S(t)\omega(t)dt. \end{cases} \tag{ComConv}$$

Here, $\mathrm{Sym}^{+,0}_{n,1}(a,b,\mathcal{T})$ is the subset of elements of $\mathrm{Sym}^{+,0}_{n}(a,b,\mathcal{T})$ with unit trace, and $\omega : [0,\mathcal{T}] \to \mathbb{S}^{n-1}$ verifies (16). We use $\mu_{\mathrm{c}}(a,b,n)$ to denote the minimal value obtained for (ComConv). The statement will then be a consequence of these two facts:

(a) (ComConv) admits minimizers;

(b) $\mu(a,b,T,n) = \mu_{\mathrm{c}}(a,b,n)$ and (Conv) admits minimizers with constant trace inputs.

Let us first provide an argument for Item (a). Fix $\omega_0 \in \mathbb{S}^{n-1}$ and consider sequences $(\mathcal{T}_l)_{l\geq 0}$ and $(S_l)_{l\geq 0}$ such that

$$\lim_{l\to+\infty} L(S_l,\omega_0) = \mu_c(a,b,T,n). \tag{17}$$

Up to subsequences, one can assume that $(\mathcal{T}_l)_{l\geq 0}$ converges to $\mathcal{T} \in [an,bn]$ and that $(S_l)_{l\geq 0}$ converges to some $S \in \mathrm{Sym}^{+,0}_{n,1}(a,b,S)$ for the weak-$*$ topology, since $\mathrm{Sym}^{+}_{n,1}$ is compact and convex. (To see that, first extend every $S_l \in Sym^{+}_{n,1}(a,b,\mathcal{T}_l)$ to an element of $\mathrm{Sym}^{+}_{n,1}(a,b,bn)$ using a common constant value in $\mathrm{Sym}^{+}_{n,1}$ and then extract a weak-$*$ convergent subsequence in $\mathrm{Sym}^{+}_{n,1}(a,b,bn)$.) It then follows that the sequence of trajectories of (16) corresponding to S_l and starting from ω_0 converges strongly, as l tends to infinty, to the trajectory ω of (16) which starts at ω_0 and corresponds to S. Since the cost L is linear with respect to S, this implies that $L(S_l,\omega_0)$ tends to $L(S,\omega_0)$ as l tends to $+\infty$, which yields that $L(S,\omega_0) = \mu_{\mathrm{c}}(a,b,n)$.

We next turn to an argument for Item (b). We set

$$\mathcal{T} := \int_0^T \mathrm{Tr}(S(t))dt, \quad \forall\, S \in \mathrm{Sym}^{+}_{n}(a,b,T).$$

By taking traces in (13), one gets that $\mathcal{T} \in [an,bn]$. Then, the change of time $\tau := \int_0^t \mathrm{Tr}(S(s))ds$, allows to define an increasing bijection $\tau : [0,\mathcal{T}] \to I$, where I is a subset of $[0,T]$ such that the set $I_0 = \{t \in I \mid S(t) = 0\}$ has zero (Lebesgue) measure and the control $\tilde{S}(\tau) := \frac{S(t)}{\mathrm{Tr}(S(t))}$ for $\tau \in [0,\mathcal{T}]$ belongs to $\mathrm{Sym}^{+}_{n,1}(a,b,\mathcal{T})$. This implies at once that $\mu(a,b,T,n) \geq \mu_{\mathrm{c}}(a,b,n)$. Pick now a minimizer $\tilde{S}$ of (ComConv), which is defined on $[0,\mathcal{T}]$ for some $\mathcal{T} \in [an,bn]$. Setting $S(t) := \mathcal{T}\tilde{S}(\mathcal{T}t/T)/T$, for $t \in [0,T]$, one has $\int_0^T S(t)dt = \int_0^{\mathcal{T}} \tilde{S}(s)ds$, and then $S \in \mathrm{Sym}^{+}_{n}(a,b,T)$ and it has constant trace. This shows that $\mu(a,b,T,n) \leq \mu_{\mathrm{c}}(a,b,n)$, proving Item (b). $\square$

The following observation will be crucial in the sequel.

Proposition 9. *The map $n \mapsto \mu(a, b, n)$ is non-increasing.*

Proof. Consider an admissible trajectory ω of (16) in dimension n, associated with some $S \in \mathrm{Sym}_n^{+,0}(a, b, T)$ and $\omega_0 \in \mathbb{S}^{n-1}$. Then, the trajectory $\tilde{\omega} = (\omega, 0)$ is a trajectory of (16) in dimension $m \geq n$ associated with $\tilde{S} = \mathrm{diag}(S, a\,\mathrm{Id}_{m-n}) \in \mathrm{Sym}_m^{+,0}(a, b, T)$ and initial condition $\tilde{\omega}_0 = (\omega_0, 0)$. Due to the form of $\tilde{\omega}$, we trivially have that $J(S, \omega_0) = J(\tilde{S}, \tilde{\omega}_0)$, and thus $\mu(a, b, n) \geq \mu(a, b, m)$ □

3. Reduction of the main results to Proposition 5

In this section, we show that Theorems 1, 3, and 4, all follow from Proposition 5. To this aim, we start by determining the homogeneity with respect to T of the quantities at hand.

Proposition 10. *For every $T > 0$ it holds*

$$R(a, b, T, n) = \frac{R(a, b, 1, n)}{T} \quad \textit{and} \quad \gamma(a, b, T, n) = T\,\gamma(a, b, 1, n) \tag{18}$$

Proof. If $c \in \mathcal{C}_n(a, b, T)$, then letting $\tilde{c}(s) := \sqrt{T}c(Ts)$, we have that $\tilde{c} \in \mathcal{C}_n(a, b, 1)$ and that $\Phi_{\tilde{c}}(s, 0) = \Phi_c(Ts, 0)$ for all $s > 0$. This immediately implies the first part of the statement. On the other hand, if $x(\cdot)$ is the trajectory of (5) associated with c and $u \in L_2((0, +\infty), \mathbb{R}^n)$, then it is also associated with $\tilde{c}$ and $\tilde{u}(s) := Tu(Ts)$. This yields at once that $\gamma(c) = T\gamma(\tilde{c})$, completing the proof. □

We are now ready to establish the link between the minimal value $\mu(a, b, n)$ of (OCP) and the worst rate of exponential decay for (DGF). We observe that this yields at once the fact that Theorem 1 is a consequence of Proposition 5.

Proposition 11. *It holds that,*

$$\frac{\mu(a, b, n)}{T} \leq R(a, b, T, n) \leq 2\frac{\mu(a/2, b/2, n)}{T} \tag{19}$$

Proof. Thanks to Proposition 10, we can restrict to the case $T = 1$. Let $(c_l)_{l\geq 0} \subset \mathcal{C}_n(a, b)$ be a minimizing sequence for $R(a, b, 1, n)$, i.e., such that there exists a vanishing sequence of positive numbers $(\varepsilon_l)_{l\geq 0}$ satisfying $R(c_l) \leq R(a, b, 1, n) + \varepsilon_l$ for $l \geq 0$. By the definition of the object at hand, there exists an increasing sequence $(t_l)_{l\geq 0}$ of times tending to infinity and a sequence $(\omega_l)_{l\geq 0}$ of unit vectors such that, for every $l \geq 0$, it holds

$$\ln \|\Phi_{c_l}(t_l, 0)\omega_l\| = \ln \|\Phi_{c_l}(t_l, 0)\| \geq (-R(c_l) - \varepsilon_l)t_l.$$

Fix $l \geq 0$ large. Set

$$k := \lfloor t_l \rfloor,\ y_0 := \Phi_{c_l}(t_l - k, 0)\omega_l,\ y_{j+1} := \Phi_{c_l}(t_l - (k-j-1), t_l - (k-j))y_j,\ 0 \leq j \leq k-1.$$

From (10), we then get

$$\ln\left(\frac{\|\Phi_{c_l}(t_l,0)\omega_l\|}{\|y_0\|}\right)=\sum_{j=0}^{k-1}\ln\left(\frac{\|y_{j+1}\|}{\|y_j\|}\right)\le -k\mu(a,b,n).$$

Clearly, there exists a positive constant $K\le 1$ independent of $l\ge 0$ such that $K\le \|y_0\|\le 1$. Thus, since $(t_l)_{l\ge 0}$ is unbounded, we deduce at once that

$$-R(a,b,1,n)-2\varepsilon_l\le -k\mu(a,b,n)/t_l.$$

By letting l tend to infinity, this yield the l.h.s. of (19).

The r.h.s. of (19) will follow from the inequality $R(2a,2b,2,n)\le\mu(a,b,n)$ to be proved next. Let $c_*\in\mathcal{C}_n(2a,2b,2)$ be the 2-periodic control given by Proposition 7 for $T=1$. It then follows from the latter and (10) that

$$\ln\|\Phi_{c_*}(k,0)\|=\sum_{\ell=1}^{k}\ln\|\Phi_{c_*}(\ell,\ell-1)\|=-k\mu(a,b,n),\qquad k\in\mathbb{N}. \tag{20}$$

Then, standard arguments yield

$$R(2a,2b,2,n)\le R(c_*)\le -\lim_{\ell\to+\infty}\frac{\ln\|\Phi_{c_*}(2\ell,0)\|}{2\ell}=\mu(a,b,n), \tag{21}$$

concluding the proof. □

We will now prove a stronger version of the statement of Theorem 3.

Theorem 12. *For every $0<a\le b$, $T>0$ and integer $n\ge 2$, one has*

$$\frac{T}{\mu(a,b,n)}\le\gamma(a,b,n,T)\le\frac{T}{1-e^{-\mu(a,b,n)}}. \tag{22}$$

Proof. Thanks to Proposition 10 it suffices to consider the case $T=1$. We start by establishing the right-hand side inequality of (22). From the variation of constant formula, for every control $u\in L_2((0,+\infty),\mathbb{R}^n)$, $c\in\mathcal{C}_n(a,b,1)$ and $t\ge 0$, the solution of (5) reads

$$x_u(t)=\int_0^t\phi_c(t,s)u(s)ds,\qquad t\ge 0. \tag{23}$$

Since it is easy to deduce from the definition of $\mu:=\mu(a,b,n)$ that $\|\phi_c(t,s)\|\le e^{-\mu\lfloor t-s\rfloor}$ for every $t\ge s\ge 0$, the above implies

$$\|x_u(t)\|\le\int_0^t e^{-\mu\lfloor t-s\rfloor}\|u(s)\|ds,\qquad t\ge 0. \tag{24}$$

Let h be the characteristic function of $\mathbb{R}_+$. Define on $\mathbb{R}$ the function $f(s)=e^{-\mu\lfloor s\rfloor}h(s)$. Then f is square integrable over $\mathbb{R}_+$ and let F be its Fourier transform. By standard computations, one has that the L_2-gain of (5) is upper bounded by $\|F\|_\infty$. It is now straightforward to observe that the supremum of F is attained at $\omega = 0$, which yields the desired upper bound.

We next give an argument for the the left-hand side inequality of (22). For that purpose, consider $c_* \in \mathcal{C}_n(a, b, 1)$ and $\omega_* \in S^{n-1}$ as provided by Proposition 5, for $T = 1$. For $t \in [0, 2]$, set

$$\rho(t) := \|\phi_{c_*}(t, 0)\omega_*\| \text{ for } t \in [0, 2], \qquad \rho_0 := \rho(2) = \exp(-2\mu) < 1. \tag{25}$$

Since c_* is 2-periodic, one has that for every $t \geq s \geq 0$ and integers $k \geq l$,

$$\phi_{c_*}(t + 2l, s + 2l) = \phi_{c_*}(t, s), \qquad \phi_{c_*}(2k, 2l)\omega_* = \rho_0^{k-l}\omega_*. \tag{26}$$

For $t \geq 0$, set $k_t := \lfloor t/2 \rfloor$ and $\xi_t = t - 2k_t$, i.e., $t = 2k_t + \xi_t$ with $\xi_t \in [0, 2)$.

For every measurable function v defined on $[0, 2]$, consider the input function $u : \mathbb{R}_+ \to \mathbb{R}^n$ given by

$$u(t) = v(\xi_t)\phi_{c_*}(\xi_t, 0)\omega_*, \qquad t \geq 0. \tag{27}$$

Observe that u is 2-periodic. Let x_u be the trajectory of (5) associated with u and starting at the origin. Then, by using (23) and (26), one has, for $t \geq 0$,

$$\begin{aligned} x_u(t) &= \int_{2k_t}^{t} \phi_{c_*}(t, s)u(s)ds + \phi_{c_*}(t, 2k_t) \sum_{j=0}^{k_t - 1} \int_{2j}^{2(j+1)} \phi_{c_*}(2k_t, s)u(s)ds \\ &= \int_{0}^{\xi_t} \phi_{c_*}(\xi_t, s)u(s)ds + \phi_{c_*}(\xi_t, 0) \sum_{j=0}^{k_t - 1} \int_{0}^{1} \phi_{c_*}(2(k_t - j), s)u(s)ds. \end{aligned}$$

Set $V(t) := \int_0^t v(s)ds$ for $t \in [0, 2]$. Thanks to (26) and (27), the above yields

$$x_u(t) = \Big(V(\xi_t) + \frac{\rho_0}{1 - \rho_0}(1 - \rho_0^{k_t})V(2)\Big)\phi_{c_*}(\xi_t, 0)\omega_*, \qquad t \geq 0. \tag{28}$$

Thus, for every positive integer k we have

$$\int_0^{2k} \|x_u(t)\|^2 dt = k \int_0^2 z^2(t)\rho^2(t)dt + O\left(\sum_{j=0}^{k-1} \rho_0^j\right),$$

where z is the function defined by

$$z(t) = V(t) + \frac{\rho_0}{1 - \rho_0}V(2), \qquad t \in [0, 2], \tag{29}$$

and the $O(\cdot)$ is uniformly bounded since $x_u(\cdot)$ is uniformly bounded. On the other hand,

$$\int_0^{2k} \|u(t)\|^2 dt = k \int_0^2 v^2(t)\rho^2(t)dt.$$

For every positive integer k, let u^k be the input function defined as follows: it is equal to u on $[0, 2k]$ and zero elsewhere. We use x^k to denote the trajectory of (5) associated with u^k and starting at the origin. Note that $x^k(t) = \phi_{c_*}(t-2k,0)x_u(2k)$, for $t \geq 2k$, which decreases exponentially to zero as t tends to infinity. Then, we have

$$\gamma(a,b,n,1) \geq \limsup_{k\to\infty} \frac{\|x^k\|_{L^2}}{\|u^k\|_{L^2}} = \sqrt{\frac{\int_0^2 z^2(t)\rho^2(t)dt}{\int_0^2 v^2(t)\rho^2(t)dt}}, \qquad \forall v \not\equiv 0. \tag{30}$$

As a consequence, thanks to the upper bound of μ given in Proposition 5, in order to complete the proof of Theorem 3, it suffices to show that there exists $v \not\equiv 0$ such that $z(t) = v(t)/\mu$ for all $t \geq 0$. By definition of z, such a function v exists if and only the following equation admits a non-zero solution,

$$\frac{\rho_0}{1-\rho_0}V(2) + V(t) = \frac{1}{\mu}V'(t), \qquad V(0) = 0. \tag{31}$$

By taking into account (25), it is easy to show that this is the case. □

Remark 13. The fact that the above implies Theorem 3 is straightforward. Indeed one deduces the lower bound in the latter by Proposition 9 and (11) and on the other hand, one gets the upper bound using the r.h.s. inequality of (19) and the estimate (3).

We finally provide the proof of Theorem 4, relying on the validity of Proposition 5.

Proof of Theorem 4. As explained in the introduction, it is enough to prove that the condition provided in the statement of the theorem is a necessary condition for GAS. Consider the three sequences $(a_l)_{l\geq 1}$, $(b_l)_{l\geq 1}$ and $(\tau_l)_{l\geq 1}$ verifying the assumptions of the theorem. For every $l \geq 1$, we define $T_l := \tau_{l+1} - \tau_l$ and apply Proposition 5 to (a_l, b_l, T_l) to deduce that there exists c_l in $\mathcal{C}_n^0(a_l, b_l, T_l)$ and $w_l \in S^{n-1}$ such that $J(c_l, w_l) = \mu(a_l, b_l, T_l)$ and the trajectory of (9) starting at w_l and corresponding to c_l is $2T$-periodic.

Choose a sequence $(U_l)_{l\geq 0}$ in $\mathrm{O}(n)$ such that, if C_l is the function defined on $[0,\tau_l]$ as the concatenation of the $U_j c_j$, $0 \leq j \leq l-1$ and if $(y_j)_{0\leq j\leq l-1}$ is the sequence defined by $y_0 := w_0$ and $y_{j+1} := \Phi_{C_l}(\tau_{j+1}, \tau_j)y_j$, then one has, for $0 \leq j \leq l-1$, that

$$\frac{\|y_{j+1}\|}{\|y_j\|} = \|\Phi_{c_j}(T_j, 0)w_0\|.$$

By summing up these relations and using the definitions of the objects at hand, one obtains

$$-\ln \|\Phi_{C_l}(\tau_l, 0)\omega_0\| = \sum_{j=0}^{l-1} \mu(a_j, b_j, n).$$

From Proposition 5, one deduces that the series of general term $\mu(a_l, b_l, n)$ converges if and only if the series of general term $\frac{a_l}{1+b_l^2}$ converges. Together with the above equation, one easily concludes. □

4. Existence of rank one periodic minimizers for (OCP)

In this section, we prove the second part of Proposition 5. This is done via the following.

Proposition 14. *There exists $c_* \in \mathcal{C}_n(a,b,T)$ and an initial condition $\omega_0 \in \mathbb{S}^{n-1}$ such that*

$$\omega_*(t) = \frac{\Phi_{c_*}(t,0)\omega_0}{\|\Phi_{c_*}(t,0)\omega_0\|}, \tag{32}$$

is $2T$-periodic and, letting $S_* = c_* c_*^\top$, both $t \mapsto S_*|_{[0,T]}(t)$ and $t \mapsto S_*|_{[T,2T]}(T-t)$, together with their respective initial conditions ω_0 and $\omega_*(T)$, are minimisers for (Conv).

In order to prove the above, we apply the Pontryagin Maximum Principle (PMP for short) to the minimizer with constant trace of (Conv) given by Proposition 7. For that purpose, we parameterize the constraint defined by (13) by introducing the variable $Q \in \mathrm{Sym}_n$ and the dynamics

$$\dot{Q}(t) = S(t), \qquad Q(0) = 0. \tag{33}$$

The PE constraint (13) now reads $Q(T) \in Sym_n(a,b)$.

The optimal control problem (Conv) can now be formulated as follows: minimize $J(S,\omega_0)$ with respect to $S \in Sym_n^{+,0}(a,b,T)$ and $\omega_0 \in \mathbb{S}^{n-1}$ along trajectories of

$$\begin{aligned} \dot{\omega} &= -S\omega + \left(\omega^T S\omega\right)\omega, && (34)\\ \dot{Q} &= S, && (35) \end{aligned}$$

starting at $(\omega_0, 0)$, and so that $Q(T) \in \mathrm{Sym}_n(a,b)$. The state space of the system is $\mathcal{M} = \mathbb{S}^{n-1} \times \mathrm{Sym}_n$. We will henceforth identify the cotangent space at $(\omega, Q) \in \mathcal{M}$ with $T^*_\omega \mathbb{S}^{n-1} \times T^*_Q \mathrm{Sym}_n \simeq (\mathbb{R}\omega)^\perp \times \mathrm{Sym}_n$.

According to the PMP, a necessary condition for solutions (ω, Q) of the optimal control problem (Conv) is the fact of being the projection of an *extremal*, i.e., an integral curve $\lambda \in T^*\mathcal{M}$ of the so-called Hamiltonian vector on $T^*\mathcal{M}$ satisfying certain additional conditions. We hereby present a definition of extremal adapted to our setting. The fact that this is equivalent to the standard definition of extremal is the subject of the subsequent proposition.

Definition 15. *A curve $\lambda : [0,T] \to T^*\mathcal{M}$ is an extremal with respect to the control $S \in \mathrm{Sym}_n^{+,0}(a,b,T)$ and $\omega_0 \in \mathbb{S}^{n-1}$ if:*

(i) letting $\lambda = (\omega, Q, p, P_Q)$, it satisfies

$$\begin{aligned} \dot{\omega} &= -S\omega + \left(\omega^T S\omega\right)\omega, && (36)\\ \dot{Q} &= S, && (37)\\ \dot{p} &= Sp - (\omega^\top S\omega)p - \dot{\omega}, && (38)\\ \dot{P}_Q &= 0. && (39) \end{aligned}$$

(ii) It holds that $p(0) = p(T) = 0$ *and that* $-P_Q$ *belongs to the normal cone of* $\mathrm{Sym}_n(a,b)$ *at* $Q(T)$*.*

(iii) Let

$$M := P_Q - (\omega p^\top + p\omega^\top + \omega\omega^\top) \quad \text{on } [0,T]. \tag{40}$$

Then, $M \le 0$ *and* $MS = SM \equiv 0$ *on* $[0,T]$*.*

We then get the following.

Proposition 16. *Let* $(\omega, Q) : [0,T] \to \mathcal{M}$ *be an optimal trajectory of the optimal control problem* (Conv)*, whose optimal control* S *has constant trace. Then* (ω, Q) *is the projection on* $\mathcal{M}$ *of an extremal* $\lambda : [0,T] \to T^*\mathcal{M}$*.*

Proof. The Hamiltonian of the system is $H(\omega, Q, p, P_Q, S) = \mathrm{Tr}(S\tilde{M})/2$, where $\tilde{M} \in \mathrm{Sym}_n$ is defined by

$$\tilde{M} = P_Q - \left(\omega p^\top + p\omega^\top + \nu_0\omega\omega^\top\right), \quad \text{with } \nu_0 \in \{0,1\}. \tag{41}$$

Recall that the existence of $(\omega, Q) : [0,T] \to \mathcal{M}$ as an optimal trajectory associated with a control S of constant trace is guaranteed by Proposition 7. After some computations, deferred to Proposition 31 in Appendix A, the PMP implies that there exists a curve $t \in [0,T] \mapsto (p(t), P_Q(t))$ and $\nu_0 \in \{0,1\}$ with $(p(t), P_Q(t), \nu_0) \neq 0$ a.e. on $[0,T]$ such that

1. $(p(t), P_Q(t)) \in T^*_{\omega(t)}\mathbb{S}^{n-1} \times T^*_{Q(t)}\mathrm{Sym}_n$ satisfy on $[0,T]$ the adjoint equations:

$$\dot{p} = Sp - (\omega^\top S\omega)p - \nu_0\dot{\omega}, \tag{42}$$
$$\dot{P}_Q = 0; \tag{43}$$

2. letting $\lambda(t) = (\omega(t), Q(t), p(t), P_Q)$ we have the maximality condition:

$$H(\lambda(t), S(t)) = \max_{S \in \mathrm{Sym}_n^+} H(\lambda(t), S) \qquad \text{a.e. on } [0,T]; \tag{44}$$

3. we have the transversality conditions:

$$p(0) \perp T^*_{\omega(0)}\mathbb{S}^{n-1}, \qquad p(T) \perp T^*_{\omega(T)}\mathbb{S}^{n-1}, \tag{45}$$

and $-P_Q$ belongs to the normal cone of $\mathrm{Sym}_n(a,b)$ at $Q(T)$.

We claim that the maximization condition implies that

$$H(\lambda(t), S(t)) \equiv 0 \quad \text{and} \quad \tilde{M} \le 0 \quad \text{on } [0,T]. \tag{46}$$

Indeed, $H(\lambda(t), 0) = 0$ and if there exists $\bar{S} \in \mathrm{Sym}_n^+$ such that $H(\lambda(t), \bar{S}) > 0$ the maximum in (44) would be infinite, since $H(\lambda(t), \gamma\bar{S}) \to +\infty$ as $\gamma \to +\infty$, proving the first part of the claim. As a consequence, $\mathrm{Tr}(S\tilde{M}) \leq 0$ for every $S \in \mathrm{Sym}_n^+$. By choosing $S = cc^\top$ for $c \in \mathbb{S}^{n-1}$, this shows that $\tilde{M} \leq 0$.

Let us now prove that $\nu_0 = 1$, which will yield Item *(i)*. We argue by contradiction and assume $\nu_0 = 0$. In this case, (42) is a linear ODE, and due to Item *(ii)*, its solution is $p \equiv 0$. This and (46) imply that $P_Q \leq 0$ and $\mathrm{Tr}(P_Q S) \equiv 0$. Since P_Q and S are symmetric real matrices which are non positive and non negative, respectively, this implies that $P_Q S \equiv 0$. Integrating over $[0, T]$ this relation yields $P_Q Q(T) = 0$, which implies $P_Q = 0$ since $Q(T) \geq a\,\mathrm{Id}$ and hence is invertible. This, however, contradicts the fact that $(p, P_Q, \nu_0) \neq 0$, thus showing that $\nu_0 = 1$.

Since for $\nu_0 = 1$ we have $M = \tilde{M}$, Item *(iii)* follows from the above and (46). On the other hand, Item *(ii)* is an immediate consequence of the transversality conditions, since $p \in T_\omega^*\mathbb{S}^{n-1}$ by definition. □

We will also need the following.

Proposition 17. *Let $\lambda = (\omega, Q, p, P_Q)$ be an extremal with respect to an optimal control S. Then, up to an orthonormal change of basis, there exists $k, r \in \mathbb{N}$, with $n = 1 + k + r$, $\alpha \in (0, 1]$, and positive definite diagonal matrices $D_Q \in \mathbb{R}^{r\times r}$ and $D_b \in \mathbb{R}^{k\times k}$, with all elements of D_Q belonging to the interval $[a, b]$, such that*

$$Q(T) = \mathrm{diag}(a, b\,\mathrm{Id}_k, D_Q) \qquad \text{and} \qquad P_Q = \mathrm{diag}(\alpha, -D_b, 0_r). \tag{47}$$

Proof. We start by showing that one can assume that P_Q and $Q(T)$ commute. To this aim, we consider a parametrized version of (Conv), which we will refer to as (P-Conv), and reads as follows: minimize $J(S, \omega_0)$ with respect to $S \in \mathrm{Sym}_n^{+,0}(a, b, T)$ and $\omega_0 \in \mathbb{S}^{n-1}$ along trajectories of

$$\begin{aligned} \dot{\omega} &= -S\omega + \left(\omega^\top S\right)\omega, &(48)\\ \dot{Q} &= USU^\top, &(49)\\ \dot{U} &= 0, &(50) \end{aligned}$$

starting at any $(\omega_0, 0, U)$ with $\omega_0 \in \mathbb{S}^{n-1}$, $U \in O(n)$, and so that $Q(T) \in \mathrm{Sym}_n(a, b)$. It is clear that (P-Conv) and (Conv) have the same minimal values and, moreover, (ω, Q, U) is a minimizer for (P-Conv) if and only if $(\omega, U^\top QU)$ is a minimizer for (Conv).

The state space of (P-Conv) is $\mathcal{M} = \mathbb{S}^{n-1} \times \mathrm{Sym}_n \times O(n)$. We denote elements of $T^*\mathcal{M}$ by $\lambda = (\omega, Q, U, p, P_Q, p_U) \in \mathbb{S}^{n-1} \times \mathrm{Sym}_n \times O(n) \times T_\omega^*\mathbb{S}^{n-1} \times T_Q^*\,\mathrm{Sym}_n \times T_U^* O(n)$. The Hamiltonian is then

$$H(\lambda) = -p^\top S\omega + \frac{\mathrm{Tr}(USU^\top P_Q)}{2} - \frac{\nu_0}{2}\omega^\top S\omega. \tag{51}$$

This shows that $(\omega, Q, U, p, P_Q, p_U)$ being an extremal for (P-Conv) is equivalent to $(\omega, U^\top QU, p, U^\top P_Q U)$ being an extremal for (Conv). In particular, in order to prove the claim, it suffices to show that $[Q(T), P_Q] = 0$ for an extremal of (P-Conv).

To this purpose, we apply the PMP to (P-Conv). As explained in Proposition 31 in Appendix A, under the identification $T^*_U O(n) \simeq U\,\mathrm{Skew}_n$, this yields the following dynamics for the costate p_U of an extremal associated with (P-Conv):

$$\dot{p}_U = U[S, U^\top P_Q U]. \tag{52}$$

Moreover, since there is no constraint on U, we have that $p_U(0) = p_U(T) = 0$. Finally, integrating (52) on $[0,T]$ yields the desired commutation relation, thus proving the claim.

In order to complete the proof of (47), let (ω, Q, p, P_Q) be an extremal for (Conv), as given by Proposition 16. Since $Q(T)$ and P_Q commute, they can be diagonalized on a common basis. Up to permuting the elements of this basis, this yields that

$$Q(T) = \mathrm{diag}(a\,\mathrm{Id}_\ell, b\,\mathrm{Id}_k, D_Q) \qquad \text{and} \qquad P_Q = \mathrm{diag}(D_a, -D_b, D), \tag{53}$$

where $n = \ell + k + r$ and D_a, D_b and D are diagonal matrices. It is then standard to show, by Proposition 16 and Item *(ii)* of Definition 15, that $D = 0$ and $D_a, D_b \geq 0$, cf. [9]. From the fact that $Q(T) \in \mathrm{Sym}_n(a,b)$ it follows that the elements of D_Q belong to the interval (a,b). Finally, up to permutations, we can redefine k, r, D_Q, and D_b in order to have $D_b > 0$, although the elements of D_Q can now belong to $[a,b]$. Thus, we are left to prove that P_Q has exactly one simple positive eigenvalue $\alpha \in (0,1]$.

Let us first show that P_Q has at most one positive eigenvalue. Indeed, by Item *(iii)* of Proposition 16, $P_Q - \omega_0\omega_0^\top$ is negative semi-definite. Therefore, the restriction of the quadratic form defined by P_Q to $(\mathbb{R}\omega_0)^\perp$ is also negative semi-definite. This implies that P_Q has at least $n-1$ non positive eigenvalues.

Thus, in order to complete the proof of (47), it suffices to contradict the fact that P_Q has no positive eigenvalues. In this case, P_Q is negative semi-definite, and therefore, one has for every $t_1 \leq t_2$ in $[0,T]$,

$$\mathrm{Tr}\Big(P_Q\big(Q(t_2) - Q(t_1)\big)\Big) \leq 0. \tag{54}$$

Let $T_0 \leq T$ be the largest time in $[0,T]$ such that $S\omega \equiv 0$ and $p \equiv 0$ on $[0,T_0]$. We first prove that T_0 exists and is strictly positive. For that purpose, pick $\bar{t} \in [0,T]$ such that $\|p(\bar{t})\| = \max\{\|p(s)\| \mid 0 \leq s \leq \bar{t}\}$ and $\mathrm{Tr}(Q(\bar{t})) \leq 1/2$. One deduces that

$$\int_0^{\bar{t}} p^\top S p\, dt \leq \max_{s\in[0,\bar{t}]} \|p(s)\|^2 \int_0^{\bar{t}} \lambda_{max}(S)\, dt \leq \|p(\bar{t})\|^2\, \mathrm{Tr}(Q(\bar{t})) \leq \frac{\|p(\bar{t})\|^2}{2}. \tag{55}$$

We next prove the following two inequalities, holding for every $t_1 \leq t_2$ in $[0,T]$,

$$\begin{aligned} \tfrac{1}{2} \quad & \left(\|p(t_2)\|^2 - \|p(t_1)\|^2\right) + \int_{t_1}^{t_2} (\omega^T(t)S(t)\omega(t))\|p(t)\|^2 dt \\ = \quad & \int_{t_1}^{t_2} p^T(t)S(t)p(t)dt + \int_{t_1}^{t_2} p^T(t)S(t)\omega^T(t)dt, \end{aligned} \tag{56}$$

and

$$\mathrm{Tr}\Big(P_Q\big(Q(t_2) - Q(t_1)\big)\Big) = 2\int_{t_1}^{t_2} p^T(t)S(t)\omega^T(t)dt + \int_{t_1}^{t_2} \omega^T(t)S(t)\omega(t)dt. \tag{57}$$

Both inequalities follow by Proposition 16: for the first one, we multiply by $p^\top$ the dynamics of $\dot p$ given by (38), integrate it on $[t_1, t_2]$, and then use the transversality conditions of Item *(ii)*. We integrate (40) over $[t_1, t_2]$ to obtain the second one.

It is immediate to deduce from (56) and (57) that, for every $t_1 \le t_2$ in $[0, T]$,

$$\int_{t_1}^{t_2} p^T(t)S(t)p(t)dt = \int_{t_1}^{t_2} \omega^T(t)S(t)\omega(t)dt + \frac{1}{2}\big(\|p(t_2)\|^2 - \|p(t_1)\|^2\big)$$
$$+ \int_{t_1}^{t_2} (\omega^T(t)S(t)\omega(t))\|p(t)\|^2 dt - \frac{1}{2}\operatorname{Tr}\Big(P_Q\big(Q(t_2) - Q(t_1)\big)\Big). \tag{58}$$

By using (54) and (58) with $t_1 = 0$ and $t_2 = \bar t$, one deduces that

$$\int_0^{\bar t} \omega^\top S\omega\, dt \le \operatorname{Tr}(P_Q Q(\bar t)) \le 0.$$

This immediately implies that $S\omega \equiv 0$, $p \equiv 0$ on $[0, \bar t]$, proving the existence of T_0 as claimed. Note that, necessarily $T_0 < T$, since otherwise one would have that $\omega \equiv \omega_0$ on $[0, T]$ and, integrating $S\omega \equiv 0$ on $[0, T]$ would yield that $Q(T)\omega_0 = 0$, contradicting the fact that $Q(T) \ge a\,\mathrm{Id}$.

We next pick $T_0 + \bar t \in [0, T]$ such that $\|p(T_0 + \bar t)\| = \max\{\|p(s)\| \mid T_0 \le s \le T_0 + \bar t\}$ and $\operatorname{Tr}(Q(T_0 + \bar t)) - \operatorname{Tr}(Q(T_0)) \le 1/2$. We then reproduce the argument starting in (55) where we replace the pair of times $(0, \bar t)$ by the pair of times $(T_0, T_0 + \bar t)$. In that way, we extend the interval on which both $S\omega$ and p are zero beyond T_0, hence contradicting the definition of T_0. We have completed the argument for the existence of a unique positive eigenvalue α for P_Q. Since $P_Q \le \omega_0\omega_0^\top$, one deduces at once that $\alpha = \lambda_{\max}(P_Q) \le \lambda_{\max}(\omega_0\omega_0^\top) = 1$. □

Proposition 18. *Let $\lambda = (\omega, Q, p, P_Q)$ be an extremal of* (Conv) *associated with a control S. Assume that, in the notations of Propositions 17, one has that $r \ge 1$. Then, there exist $(\tilde\omega, \tilde p) \in T^*\mathbb{S}^k$, $\tilde S \in \mathrm{Sym}^{+,0}_{k+1}(a, b, T)$ and $S_0 \in \mathrm{Sym}^{+,0}_r(a, b, T)$, such that*

$$\omega = (\tilde\omega, 0)^\top, \qquad p = (\tilde p, 0)^\top \qquad \text{and} \qquad S = \operatorname{diag}(\tilde S, S_0). \tag{59}$$

Moreover, letting $\tilde Q = \operatorname{diag}(a, b\,\mathrm{Id}_k)$ and $\tilde P_Q = \operatorname{diag}(\alpha, -D_b)$, we have that $\tilde\lambda = (\tilde\omega, \tilde Q, \tilde p, \tilde P_Q)$ is an extremal trajectory with control $\tilde S$ of (Conv) *in dimension $k+1$, and $J(\tilde S, \tilde\omega(0)) = J(S, \omega(0))$. In particular, if $k = 0$, this implies $p \equiv 0$, $\omega \equiv \omega_0 \in \mathbb{S}^{n-1}$, and $J(S, \omega_0) = a$.*

Proof. We start by decomposing $\omega = (\tilde\omega, \omega_0)$ and $p = (\tilde p, p_0)$ for some $\omega_0, p_0 \in \mathbb{R}^r$. Our aim is to prove that $p_0 \equiv \omega_0 \equiv 0$. Let us define

$$\tilde A = (\tilde p + \tilde\omega)(\tilde p + \tilde\omega)^\top - \tilde p\tilde p^\top, \qquad A_0 = (p_0 + \omega_0)(p_0 + \omega_0)^\top - p_0 p_0^\top \tag{60}$$
$$B = (\tilde p + \tilde\omega)(p_0 + \omega_0)^\top - \tilde p p_0^\top \tag{61}$$

Then, by Item *iii.* of Proposition 16, we get

$$\begin{pmatrix} \tilde A - \tilde P_Q & B \\ B^\top & A_0 \end{pmatrix} \ge 0. \tag{62}$$

We deduce at once that $A_0 \geq 0$, and thus, that there exists $\varrho \in [-1,1]$ such that $p_0 = \varrho(p_0 + \omega_0)$. In particular, $p_0 + \omega_0 = 0$ if and only if $p_0 = \omega_0 = 0$. Let I be a maximal open interval such that $p_0 + \omega_0 \neq 0$ and assume, by contradiction, that $I \neq \varnothing$.

We claim that

$$-(1-\varrho^2)\tilde{P}_Q \geq \big((1-\varrho)\tilde{p} + \varrho\tilde{\omega}\big)\big((1-\varrho)\tilde{p} + \varrho\tilde{\omega}\big)^\top \qquad \text{on } I. \tag{63}$$

To this effect, set $A_\varepsilon = A_0 + \varepsilon(p_0 + \omega_0)(p_0 + \omega_0)^\top$ for $\varepsilon > 0$. Observe that (62) holds with A_0 replaced by A_ε. Then, by Schur complement formula we have

$$\tilde{A} - \tilde{P}_Q - BA_\varepsilon^\dagger B^\top \geq 0, \tag{64}$$

where we denoted by $A_\varepsilon^\dagger$ the Moore-Penrose inverse of A_ε. Let us observe that

$$A_\varepsilon = (1-\varrho^2+\varepsilon)(p_0+\omega_0)(p_0+\omega_0)^\top \quad \text{and} \quad B = ((1-\varrho)\tilde{p} + \tilde{\omega})\,(p_0+\omega_0)^\top. \tag{65}$$

Since $A_0^\dagger = \frac{(p_0+\omega_0)(p_0+\omega_0)^\top}{(1-\varrho^2+\varepsilon)\|p_0+\omega_0\|^4}$, the claim follows by letting $\varepsilon \downarrow 0$ in (64) and simple computations.

In order to obtain the desired contradiction, we observe that it has to hold $\varrho^2 = 1$. Indeed, by (63), we have $-(1-\varrho^2)\alpha \geq 0$ with $\alpha > 0$. On the other hand, if $\varrho = 1$, we have $\omega_0 \equiv 0$ on I by definition of ϱ, and $\tilde{\omega} \equiv 0$ on I by (63), which contradicts $\omega \in \mathbb{S}^{n-1}$. Thus, $\varrho = -1$ and thus, by (63), it holds $2p + \omega \equiv 0$ on I. However, since $p \in \omega^\perp$, we have $\|2p + \omega\| \geq \|\omega\| \equiv 1$, thus yielding the desired contradiction. This implies that $I = \varnothing$, and thus that $\omega_0 \equiv p_0 \equiv 0$ on $[0,T]$.

Finally, a straightforward computation yields the desired form for S, together with the fact that $\tilde{\lambda}$ is an extremal trajectory with control $\tilde{S}$ of (Conv) in dimension $k+1$. The last part of the statement follows by the explicit computation of the solutions in dimension $n = 1$. □

As we will see, the above Proposition immediately yields Proposition 14 if we are in the case $k = 0$. Thus, we henceforth focus on extremals that satisfy the following assumption:

Assumption 1. The extremal $\lambda = (\omega, Q, p, P_Q)$ is such that

$$Q(T) = \operatorname{diag}(a, b\,\mathrm{Id}_{n-1}) \qquad \text{and} \qquad P_Q = \operatorname{diag}(\alpha, -D_b), \tag{66}$$

where $\alpha \in (0,1]$, and $D_b \in \mathbb{R}^{(n-1)\times(n-1)}$ is a positive definite diagonal matrix.

We start by proving some essential properties of the matrix

$$M = P_Q - \Big(\omega p^\top + p\omega^\top + \omega\omega^\top\Big). \tag{67}$$

We recall that, by Item *iii.* of Proposition 16, we have

$$M \leq 0 \qquad \text{and} \qquad MS \equiv SM \equiv 0. \tag{68}$$

Proposition 19. *Let λ be an extremal satisfying Assumption 1. Then,* $\operatorname{rank} M \equiv n-1$*, and M has constant spectrum (taking into account multiplicities). In particular,* $\operatorname{rank} S \equiv 1$.

Proof. By (40), we have $\operatorname{rank} M \le n-1$. Let $x \in \ker M$, $x \neq 0$, then

$$x = \big((p+\omega)^\top x\big) P_Q^{-1}\omega + (\omega^\top x) P_Q^{-1} p. \tag{69}$$

As a consequence, it holds $\ker M \subset V := \operatorname{span}\{P_Q^{-1}\omega, P_Q^{-1}p\}$. Observe that $\dim V \in \{1,2\}$, with $\dim V = 2$ if and only if $p \neq 0$. In particular, $\operatorname{rank} M \ge n-2$.

Let $E = \{t \in [0,T] \mid \operatorname{rank} M(t) = n-1\}$. Trivially, E is open in $[0,T]$, and, moreover, $E \neq \varnothing$, since $\{0,T\} \subset E$. We now show that E is also closed, which implies $E = [0,T]$ thus completing the proof of the statement.

To this aim, let us start by observing that, thanks to (68), on E it holds $S = cc^\top$ for some $c \in \ker M$. Since, on E, an element of $\ker M$ can be expressed as an analytic function of $(\omega(t), p(t))$ by (69), the dynamics (36)-(38) imply that the extremal trajectory is locally analytic at $t \in E$. In particular, $t \mapsto M(t)$ is analytic on E. As a consequence, the (unordered) negative eigenvalues $(-\lambda_j)_{j=1}^{n-1}$ of M are analytic on E, and it is possible to find an analytic family $\{v_1, \dots, v_{n-1}\}$ of associated orthonormal eigenvectors, (see, e.g., [11, Theorem 6.1 and Section 6.2]).

Differentiating with respect to $t \in E$ the relation $v_\ell^\top M v_\ell = \lambda_\ell$ for $\ell = 1, \dots, n-1$, we have

$$v_\ell^\top \dot M v_\ell = \dot\lambda_\ell \quad \text{on } E. \tag{70}$$

To complete the proof of the statement, observe that straightforward computations from the definition of M yield

$$\dot M = -(\omega^\top c)(cp^\top + pc^\top) + (p^\top c)(c\omega^\top + \omega c^\top) \quad \text{on } E. \tag{71}$$

Using this and the fact that $c^\top v_\ell \equiv 0$ in (94) yields that $\dot\lambda_\ell \equiv 0$, $\ell = 1, \dots, n-1$. This yields at once that E is closed, completing the proof of the statement. □

Proposition 20. *Let λ be an extremal satisfying Assumption 1. Then,*

$$\omega_i(0)^2 = \omega_i(T)^2 \qquad \text{for} \qquad i = 1, \dots, n \tag{72}$$

Proof. Let us write $P_Q = \operatorname{diag}(-d_1, \dots, -d_n)$, where $d_1 = \alpha$ and $0 < d_i \le d_{i+1}$, for $2 \le i \le n$. By (68) and Proposition 19, the control associated with λ takes the form $S = cc^\top$ for some $c \in \mathcal{C}_n^0(a,b,T)$, $c \not\equiv 0$. Set

$$q = p + \omega/2, \;\; \gamma = c^\top\omega, \;\; \delta = c^\top q.$$

Then, the fact that $Mc \equiv 0$ yields

$$c_i = -{d_i}^{-1}\left((c^\top q)\omega_i + (c^\top\omega) q_i\right), \qquad i = 1, \dots, n. \tag{73}$$

Moreover, by Proposition 17, $Q = \operatorname{diag}(a, b \operatorname{Id}_{n-1})$. This implies that

$$\int_0^T c_1 c_i \, dt = a\delta_{1i} \quad \text{and} \quad \int_0^T c_i c_j \, dt = b\delta_{ij}, \quad i, j = 2, \ldots, n. \tag{74}$$

Finally, letting $Z_i = (\omega_i, q_i)^\top$ for $i = 1, \ldots, n$, we have the following dynamics:

$$\dot{Z}_i = A_i Z_i, \quad \text{where } A_i(t) := \frac{1}{d_i} \begin{pmatrix} \gamma(\delta + d_i\gamma) & \gamma^2 \\ \delta^2 & -\gamma(\delta + d_i\gamma) \end{pmatrix}. \tag{75}$$

Note that A_i depends on i only through the d_i's. The above implies at once that both $\omega_i(0), \omega_i(T)$ are non zero since, otherwise, from $p_i(0) = p_i(T) = 0$, we would have $Z_i(0)$ or $Z_i(T) = 0$, and thus $Z_i \equiv 0$. However, this would yield $c_i \equiv 0$, contradicting (74).

We now claim that the d_i's are two by two distinct. Indeed, if this were not the case, we would have $d_i = d_{i+1}$ for some $i \geq 2$. By (75), this implies that

$$Z_{i+1} = \frac{\omega_{i+1}(0)}{\omega_i(0)} Z_i \quad \text{on } [0, T]. \tag{76}$$

Then, by (73), we have

$$c_{i+1} = \frac{\omega_{i+1}(0)}{\omega_i(0)} c_i \quad \text{on } [0, T], \tag{77}$$

which yields

$$\int_0^T c_i c_{i+1} \, dt = \frac{\omega_{i+1}(0)}{\omega_i(0)} \int_0^T c_i^2 \, dt = b \frac{\omega_{i+1}(0)}{\omega_i(0)} \neq 0. \tag{78}$$

However, this contradicts (74) and thus proves the claim.

Let us now denote by $\mathfrak{p}$ and $\mathfrak{m}$ the characteristic polynomials of the matrices P_Q and M, respectively. Observe that, since P_Q and $\operatorname{Spec}(M)$ (taking into account multiplicities) are independent of $t \in [0, T]$, the same is true for $\mathfrak{p}$ and $\mathfrak{m}$. In order to complete the proof of the statement, we will compute the ratio $\mathfrak{m}/\mathfrak{p}$ in two different ways.

Firstly, we observe that, by definition of M, it holds

$$\frac{\mathfrak{m}(\xi)}{\mathfrak{p}(\xi)} = \det\Big(\operatorname{Id}_n + (\xi \operatorname{Id}_n - P_Q)^{-1}(\omega\omega^\top + p\omega^\top\omega p^\top)\Big), \qquad \xi \in \mathbb{R}. \tag{79}$$

A simple computation yields $(\xi \operatorname{Id}_n - P_Q)^{-1}(\omega\omega^\top + p\omega^\top\omega p^\top) = vw^\top$, where

$$v = (\xi \operatorname{Id}_n - P_Q)^{-1}\omega, (\xi \operatorname{Id}_n - P_Q)^{-1}p) \qquad \text{and} \qquad w = (\omega + p, \omega). \tag{80}$$

Then, using Sylvester's determinant identity we have, for $\xi \in \mathbb{R}$,

$$\frac{\mathfrak{m}(\xi)}{\mathfrak{p}(\xi)} = \det(\operatorname{Id}_2 + w^\top v) = 1 + \sum_{i=1}^n \left(\omega_i^2 + 2\omega_i p_i + \sum_{j \neq i} \frac{(\omega_i p_j - \omega_j p_i)^2}{d_i - d_j} \right) \frac{1}{\xi + d_i}. \tag{81}$$

We have used, in the last equality, the fact that the d_i's are two by two distinct.

On the other hand, by decomposing the rational fraction $\mathfrak{m}/\mathfrak{p}$ in simple elements in terms of the indeterminate ξ, we have

$$\frac{\mathfrak{m}(\xi)}{\mathfrak{p}(\xi)} = 1 + \sum_{i=1}^{n} \frac{\mathfrak{m}(-d_i)}{\mathfrak{p}'(-d_i)} \frac{1}{\xi + d_i}. \tag{82}$$

By comparing the above with (81), we finally obtain the following integrals of motion

$$\frac{\mathfrak{m}(-d_i)}{\mathfrak{p}'(-d_i)} = \omega_i^2 + 2\omega_i p_i + \sum_{j\neq i} \frac{(\omega_i p_j - \omega_j p_i)^2}{d_i - d_j}, \qquad \text{on } [0,T]. \tag{83}$$

The statement then follows by evaluating the above at $t = 0$ and $t = T$, and using the transversality conditions $p(0) = p(T) = 0$. □

We are finally in a position to prove the main result of this section.

Proof of Proposition 14. Let $\lambda = (\omega, Q, p, P_Q)$ be an extremal of (Conv) associated with an optimal control S. By Proposition 17, up to an orthonormal change of basis, there exist $\alpha \in (0, 1]$, $r \in \mathbb{N}$, and a positive diagonal matrix $D_Q \in \mathbb{R}^{r\times r}$ such that we have the following dichotomy:

1. $r = n - 1$, $Q(T) = \operatorname{diag}(a, D_Q)$ and $P_Q = \operatorname{diag}(\alpha, 0_{n-1})$;
2. $r = n - 1 - k$ for some $k \in \{1, \ldots, n - 1\}$, $Q(T) = \operatorname{diag}(a, b\,\mathrm{Id}_k, D_Q)$ and $P_Q = \operatorname{diag}(\alpha, -D_b, 0_r)$ for some positive diagonal matrix $D_b \in \mathbb{R}^{k\times k}$;

In the first case, by Proposition 18, we have $J(S, \omega(0)) = a$. Then, it suffices to consider a periodic version $c_* \in \mathcal{C}_n(a, b, T)$ of the control $c \in \mathcal{C}_n^0(a, b, T)$ defined in Proposition 6. Indeed, it holds $J(c, \omega(0)) = a$, and thus $cc^\top$ is an optimal control, and moreover the corresponding trajectory is periodic since $\omega \equiv \omega(0)$.

Let us now focus on the second case. We start by assuming that $r = 0$, i.e., that λ satisfies Assumption 1. Then, $S = cc^\top$ for some $c \in \mathcal{C}_n^0(a, b, T)$ since $MS \equiv 0$ by (68) and $\operatorname{rank} M = n - 1$ by Proposition 19. Moreover, by Proposition 20, there exists a diagonal matrix D with entries ± 1 such that $\omega(T) = D\omega(0)$ and $c(T) = Dc(0)$. We next define the required $c_* \in \mathcal{C}_n(a, b, T)$ as follows. On $[0, T]$, it is equal to c and, for $t \in [T, 2T]$, we take $c_*(t) := Dc(t - T)$. Clearly, the corresponding trajectory ω_* starting at $\omega(0)$ will satisfy $\omega_*(t) = D\omega(t-T)$ for $t \in [T, 2T]$ due to the fact that $D \in \mathrm{O}(n)$ and the invariance of the dynamics by elements of $\mathrm{O}(n)$. In particular, $\omega_*(2T) = \omega_*(0)$ since $D^2 = \mathrm{Id}_n$ and similarly $c_*(2T) = c_*(0)$. Extending c_* on $[0, \infty)$ by $2T$-periodicity yields the result. Finally, the case $r \geq 1$ follows from the case $r = 1$, thanks to Proposition 18 and to the fact that, by the proof of Proposition 9, we can lift every trajectory in dimension $k + 1$ to a trajectory in dimension n with the same cost. □

5. The 2D case

In this section, we prove the first part of Proposition 5, which, thanks to Proposition 9 reduces to the following.

Proposition 21. *There exists an universal constant $C_0 > 0$ such that, for every $0 < a \leq b$, one has*

$$\mu(a, b, 2) \leq \frac{C_0 a}{1 + b^2}. \tag{84}$$

We start by introducing adapted notations for the 2D case. Indeed, in this case, $\mathcal{M} = \mathbb{S}^1 \times \mathrm{Sym}_2$ and, thanks to Proposition 17, up to an orthonormal change of basis, we can represent an extremal $\lambda = (\omega, Q, p, P_Q) \in T^*\mathcal{M}$ as $\lambda = (\theta, Q, \eta, \alpha, d) \in \mathbb{S}^1 \times \mathrm{Sym}_2 \times \mathbb{R} \times (0, 1] \times [0, +\infty)$, via the identifications

$$\omega = e^{i\theta/2}, \qquad p = \eta i e^{i\theta/2}, \qquad P_Q = \mathrm{diag}(\alpha, -d), \qquad Q(T) = \mathrm{diag}(a, b). \tag{85}$$

5.1. Structure of optimal trajectories

In this subsection, we consider a fixed extremal $\lambda = (\theta, Q, \eta, \alpha, d)$ satisfying Assumption 1, and associated with an optimal control $S = cc^\top$ of constant trace. The (PE) condition and (85) imply that, up to a time reparametrization, we have

$$\|c\| = 1 \qquad \text{and} \qquad T = a + b. \tag{86}$$

Moreover, the matrix M defined in (40) can be written on $[0, T]$ as

$$M = \mathrm{diag}(\alpha, -d) - p\left(\omega(\omega^\perp)^\top + \omega^\perp \omega^\top\right) - \omega\omega^\top. \tag{87}$$

Since $Mc = 0$ on $[0, T]$ and the trace of M is constant and equal to $\alpha - d - 1$, it holds

$$M(t) = (\alpha - d - 1)c^\perp(t)(c^\perp(t))^T, \qquad \text{for a.e. } t \in [0, T]. \tag{88}$$

Since this insures that c is actually absolutely continuous, this equality holds on the whole interval $[0, T]$.

In the following proposition, we rewrite the dynamics of an extremal trajectory with the adapted notations for the 2D case.

Proposition 22. *Letting $c = e^{i\phi/2}$, $\phi \in \mathbb{S}^1$, we have the following dynamics*

$$\dot{\theta} = s_{\theta - \phi}, \qquad \dot{\eta} = -\frac{s_{\theta-\phi} + 2\eta c_{\theta-\phi}}{2}, \qquad \dot{\phi} = \frac{2\eta}{1 - \alpha + d}. \tag{Σ}$$

Moreover, $\eta(0) = \eta(T) = 0$ and, for every $t_\star \in [0, T]$ such that $\eta(t_\star) = 0$, we have

$$c_{\theta(t_\star)} = 1 - \frac{2d(1 - \alpha)}{\alpha + d}, \tag{89}$$

$$c_{\phi(t_\star)} = -1 + \frac{2d(1 + d)}{\alpha - \alpha^2 + d + d^2} \tag{90}$$

In addition, the following relations hold on $[0,T]$

$$c_\theta - 2\eta s_\theta = \frac{2\eta^2 + 2\alpha d + \alpha - d}{\alpha + d}, \tag{91}$$

$$\eta(2\dot\eta - (\alpha+d)s_\phi) = 0. \tag{92}$$

Proof. The transversality conditions (Item *ii.* of Definition 15) imply immediately that $\eta(0) = \eta(T) = 0$. The first two equations of (Σ) follow at once from (36)-(38) and (85). Let us prove the last one. Due to the fact that $2\dot c = \dot\phi c^\perp$ and $2\dot c^\perp = -\dot\phi c$, differentiating (88) yields

$$\dot M = \frac{1-\alpha+d}{2}\dot\phi(c^\perp c^\top + c(c^\perp)^\top). \tag{93}$$

On the other hand, by using (85) in the expression of $\dot M$ given in (71), a standard computation yields

$$\dot M = \eta(c^\perp c^\top + c(c^\perp)^\top). \tag{94}$$

Putting together this and (93) proves the last equation of (Σ).

Equations (89) and (90) follow at once by developing the equation $M(t_\star)c(t_\star) = 0$ at every $t_\star \in [0,T]$ such that $\eta(t_\star) = 0$. Finally observe that, letting $R_{\theta/2}$ be the matrix corresponding to a clock-wise rotation by $\theta/2$, we have that

$$0 = \det M = \det\left(P_Q R_{\theta/2} - R_{\theta/2}\begin{pmatrix} 1 & \eta \\ \eta & 0\end{pmatrix}\right). \tag{95}$$

Direct computations show that the above equation is equivalent to (91). On the other hand, taking the derivative with respect to time of the above yields (92). ☐

Observe that, up to rotating ω and c, we can always assume that $\phi_0 = \phi(0) \in (0,\pi)$. We next show that the dynamic of the control c is actually independent of θ and η.

Lemma 23. *The control* $c = e^{i\phi/2}$*,* $\phi \in \mathbb{S}^1$*, satisfies the inverted pendulum equation*

$$\ddot\phi = \frac{1}{2\nu^2}s_\phi, \qquad \textit{where} \qquad \nu = \sqrt{\frac{1-\alpha+d}{2(\alpha+d)}}. \tag{96}$$

Moreover, $\left(\frac{1-\alpha+d}{2}\right)\dot\phi^2$ *takes values in* $[0, d(d+1)]$ *and there exist* $\kappa \in \mathbb{N}^*$ *such that* $\eta(t) = 0$ *if and only if* $t = jT/\kappa$*, for* $j \in [\![0,\kappa]\!]$*, and such that the following relations hold*

$$a = \nu\kappa K_+(\phi_0), \qquad b = \nu\kappa K_-(\phi_0), \quad \textit{where} \quad K_\pm(\gamma) = \int_\gamma^\pi \frac{1 \pm c_\phi}{\sqrt{c_\gamma - c_\phi}}\, d\phi. \tag{97}$$

Proof. Let $\mathcal{P} = \{t \in [0,T] \mid \eta(t) = 0\}$. Notice that the times 0 and T belong to $\mathcal{P}$ and we claim that $\mathcal{P}$ is made of a finite number of points. To see that, we argue by contradiction and hence assume that $\mathcal{P}$ is not finite. Then, since η is differentiable a.e. on $[0,T]$, there would exist $t_0 \in [0,T]$ such that $\eta(t_0) = \dot\eta(t_0) = 0$. However, this would imply that $(\theta, \eta, \phi) \equiv (0,0,0)$ is the only solution to (Σ) of Proposition 22. In particular, the control c is constant, yielding that $Q(T)$ has rank one, which contradicts the fact that $c \in \mathcal{C}_2^0(a,b,T)$. The claim that $\mathcal{P}$ is made of a finite number of points is established.

We next observe that, due to the above claim and the continuity of $\dot\eta$ given by (Σ) of Proposition 22, the formula (92) reduces to $2\dot\eta = (\alpha + d)s_\phi$ on $[0,T]$. By taking its derivative with respect to time, this yields (96).

Let us provide an argument for (97). Since $\dot Q = cc^\top$ and $Q(T) = \operatorname{diag}(a,b)$, we have $Q(T) = \int_0^T cc^\top\, dt$. By simple computations, we have that

$$cc^\top = \frac{1}{2}\left(\mathrm{Id}_2 + \begin{pmatrix} c_\phi & s_\phi \\ s_\phi & -c_\phi \end{pmatrix}\right). \tag{98}$$

This yields at once that

$$a = \frac{1}{2}\left(T + \int_0^T c_\phi\, dt\right) \qquad \text{and} \qquad b = \frac{1}{2}\left(T - \int_0^T c_\phi\, dt\right). \tag{99}$$

Let $\mathcal{P} = \{0 = t_0 < t_1 < \ldots < t_\kappa = T\}$ for some $\kappa \in \mathbb{N}^*$. We now claim that, for $j = 0, \ldots, \kappa - 1$, we have

$$t_{j+1} - t_j = 2\nu \int_{\phi_0}^{\pi} \frac{d\phi}{\sqrt{c_{\phi_0} - c_\phi}} \qquad \text{and} \qquad \int_{t_j}^{t_{j+1}} c_\phi\, dt = 2\nu \int_{\phi_0}^{\pi} \frac{c_\phi\, d\phi}{\sqrt{c_{\phi_0} - c_\phi}}. \tag{100}$$

Elementary computations show that (97) will follow. In order to prove the claim, observe that by (96), for every $j \in \{0, \ldots, \kappa - 1\}$, we have the following first integral:

$$\dot\phi^2 = \frac{c_{\phi(t_j)} - c_\phi}{\nu^2} \qquad \text{on } [0,T]. \tag{101}$$

This implies at once that $c_{\phi(t_j)}$ is independent of j, and thus $c_{\phi(t_j)} = c_{\phi_0}$. Since η has constant sign on $I_j = (t_j, t_{j+1})$, ψ is monotone and we have that s_ϕ must take the value zero at a unique $\tau_j \in I_j$. Moreover, by (101) we have $c_{\phi(\tau_j)} < c_{\phi_0}$. Since we assumed $\phi_0 \in (0,\pi)$, this implies that $\phi(\tau_j) = \pi$ and that $\phi(t) \neq \pi$ if $t \in I_j \setminus \{\tau_j\}$.

By the classical structure of the phase portrait of the inverted pendulum equation (96), see [5], we have that $\tau_j = (t_{j+1}+t_j)/2$, that $\phi(2\tau_j - t) = 2\pi - \phi(t)$ for $t \in [t_j, \tau_j]$, and that ϕ is a monotone bijection from $[t_j, \tau_j]$ to $[\phi(t_j), \pi]$ or $[\pi, \phi(t_j)]$. A simple computation shows that $\left(\frac{1-\alpha+d}{2}\right)\dot\phi^2$ takes values in $[0, d(d+1)]$ and the maximal value $d(d+1)$ is exactly at $t = \tau_j$, for every $j \in \{0, \ldots, \kappa - 1\}$.

By (101) we get (100), thus completing the proof. $\square$

5.2. Existence of extremal trajectories

Up to now, we have only provided necessary conditions on the structure of optimal trajectories. In this subsection, we show that extremal trajectories satisfying these conditions do exist. The main result is the following.

Proposition 24. *There exists $b_0 > 0$ such that, for every $0 < a \le b$ with $b \ge b_0$, there exists a control $c = e^{i\phi/2} \in \mathcal{C}_2^0(a, b, a+b)$, $(\alpha, d) \in (0,1) \times \mathbb{R}_+^*$ for which $\lambda = (\theta, \mathrm{diag}(a,b), \eta, \alpha, d)$ is an extremal satisfying Assumption 1 associated with the control $S = cc^\top$ such that the conclusions of Proposition 22 and Lemma 23 hold true, with only half of a period (i.e., $\kappa = 1$ and $\eta \neq 0$ on $(0, a+b)$). In addition, one can choose the above parameters, as b tends to infinity, in such a way that:*

(*a*) *ϕ_0 is the largest solution in $(0,\pi)$ of*

$$\frac{K_+(\phi_0)}{K_-(\phi_0)} = \frac{a}{b}, \tag{102}$$

and there exists two positive constants c_0, c_1 independent of ϕ_0 such that,

$$c_0 \frac{a}{b} \le c_{\phi_0/2}^2 \le c_1 \frac{a}{b}; \tag{103}$$

(*b*) *the following hold true,*

$$\alpha \sim_{b\to\infty} \frac{K_-(\phi_0)^2}{2b^2}, \quad \textit{and} \quad d = \frac{1 + c_{\phi_0}}{2}\alpha. \tag{104}$$

Remark 25. Note that α is always larger than d. Moreover, the above implies that there exists $c_1, c_2 > 0$ and $b_1 > 0$ such that

$$c_1 \frac{a}{b^2} \le \frac{d}{\sqrt{\alpha}} \le c_2 \frac{a}{b^2}, \qquad \forall\, 0 < a \le b \quad \text{s.t.} \quad b > b_1. \tag{105}$$

In order to prove the previous proposition, we need two technical results. In the first one, of which we omit the proof, we collect the output of simple computations.

Lemma 26. (*a*) *Let us define the following function*

$$F(\alpha, d, \xi) = \frac{2\xi + 2\alpha d + \alpha - d}{(\alpha + d)\sqrt{1 + 4\xi}}, \qquad (\alpha, d, \xi) \in (0,1) \times \mathbb{R}_+^* \times [0, d(d+1)]. \tag{106}$$

Then, for every $(\alpha, d) \in (0,1) \times \mathbb{R}_+^*$ and $\xi \in [0, d(d+1))$, one has $|F(\alpha, d, \xi)| < 1$ and $F(\alpha, d, d(d+1)) = 1$.

(*b*) *For $\gamma \in (0,\pi)$, consider the elliptic integrals functions defined previously as*

$$K_+(\gamma) = \int_\gamma^\pi \frac{1 + c_\phi}{\sqrt{c_\gamma - c_\phi}}\, d\phi, \quad K_-(\gamma) = \int_\gamma^\pi \frac{1 - c_\phi}{\sqrt{c_\gamma - c_\phi}}\, d\phi.$$

Then, K_+ is a smooth strictly decreasing bijection from $(0,\pi)$ to $(0,\infty)$, while K_- is a smooth function taking values in $[K_-(0), K_-(\pi)]$, where

$$K_+(\gamma) \sim_{\gamma\to\pi^-} \frac{\pi}{4\sqrt{2}}(\pi - \gamma)^2, \tag{107}$$

$$K_-(\pi) := \lim_{\gamma\to\pi^-} K_-(\gamma) = \sqrt{2}\pi, \qquad \textit{and} \qquad K_-(0) := \lim_{\gamma\to 0^+} K_-(\gamma) = 2\sqrt{2}. \tag{108}$$

The next lemma allows us to construct extremals corresponding to fixed (α, d).

Lemma 27. *Let $(\alpha, d) \in (0,1) \times \mathbb{R}_+^*$. Then, there exists $0 < a \leq b$, and a control $c = e^{i\phi/2} \in \mathcal{C}_2^0(a, b, a+b)$ such that if (θ, η) are given by (Σ) with initial conditions $(\theta_0, 0)$, where θ_0 verifies (89), then $\lambda = (\theta, \mathrm{diag}(a,b), \eta, \alpha, d)$ is an extremal satisfying Assumption 1 associated with the control $S = cc^\top$. Moreover, $\eta(t) \neq 0$ for $t \in (0, a+b)$. In particular, we have that*

1. *$\phi_0 = \phi(0) \in (0, \pi)$ and verifies (90);*
2. *it holds that $a = \nu K_+(\phi_0)$ and $b = \nu K_-(\phi_0)$.*

Proof. We start by observing that, by (106), equation (91) is equivalent to

$$s_{\theta_\eta - \theta} = F(\alpha, d, \eta^2), \tag{109}$$

where θ_η is the unique angle in $[0, 2\pi)$ such that $s_{\theta_\eta} = (1+4\eta^2)^{-1/2}$ and $c_{\theta_\eta} = 2\eta(1+4\eta^2)^{-1/2}$. Also note that $\frac{\partial \theta_\eta}{\partial \eta} = -\frac{2}{1+4\eta^2}$. In particular, for every $\eta_\star$ such that $c_{\theta_\eta - \theta} \neq 0$, i.e., $|F(\alpha, d, \eta_\star^2)| < 1$, the implicit function theorem allows us to write $\theta = f(\eta)$ in a neighbourhood of $\eta_\star$ where f is smooth. By taking into account Item (a) of Lemma 26, this holds as long as $\eta_\star^2 < d(d+1)$. Since $\eta(0) = 0$ along an extremal satisfying Assumption 1 and by (Σ), ϕ has to satisfy

$$c_{\theta_\eta - \theta} \left[\frac{\partial \theta_\eta}{\partial \eta} \dot{\eta} - \dot{\theta} \right] = 2 \frac{\partial F}{\partial \xi} \eta \dot{\eta}. \tag{110}$$

By (Σ) and the expression of $\partial \theta_\eta / \partial \eta$, it is then sufficient to solve the following

$$s_{\theta - \phi} \left(\frac{4\eta}{1+4\eta^2} c_{\theta_\eta - \theta} + \frac{\partial F}{\partial \xi} \right) + c_{\theta - \phi} \left(2\eta \frac{\partial F}{\partial \xi} - c_{\theta_\eta - \theta} \right) = 0. \tag{111}$$

This is clearly possible whenever $\eta^2 < d(d+1)$, and thus $\theta - \phi$ is well-defined modulo π.

To show that such a choice of a control ϕ indeed defines an extremal of the problem, it remains to prove that $Mc = 0$ as long as the control is defined, where M has been defined in (40) and $c = e^{i\phi/2}$. Along a trajectory (θ, η, ϕ) of (Σ), one has that (94) holds true. On the other hand, we let ψ be an absolutely continuous parametrization in $\mathbb{S}^1$ of $Ker\ M$, which is possible since the latter matrix is always of rank one, and set $\tilde{c} := e^{i\psi/2}$. Hence, one has

$$\dot{M} = \frac{1 - \alpha + d}{2} \dot{\psi} (\tilde{c}^\perp \tilde{c}^\top + \tilde{c} (\tilde{c}^\perp)^\top). \tag{112}$$

Comparing the latter with (94) yields

$$\tilde{c}^\perp \tilde{c}^\top + \tilde{c} (\tilde{c}^\perp)^\top = c^\perp c^\top + c (c^\perp)^\top.$$

This, implies that c and $\tilde{c}$ differ by a constant angle, which is necessarily an integer multiple of $\pi/2$, and hence $\phi - \psi = 0$ modulo π. Choosing $\phi(0) \in (0, \pi)$ verifying (90), we deduce that $Mc = 0$ as long as it is defined. It is clear that ϕ verifies the inverted pendulum equation (96) on the interval of time $[0, \tau_1)$, where τ_1 is the time where $\phi = \pi$ and $\dot{\phi} = \sqrt{d(d+1)}$. We then extend ϕ and the whole extremal trajectory on $[0, T]$, where $T := 2\tau_1$, by using the symmetry properties of the solution of (96). One can finally check, as in the proof of Lemma 23, that if we let $0 < a \leq b$ such that (97) is satisfied with $\kappa = 1$, then $c \in \mathcal{C}_2^0(a, b, a+b)$. It is then straightforward to verify that all the requirements of the statement have been established. $\square$

We are finally ready to prove the main result of this section.

Proof of Proposition 24. By Lemma 27, the proof reduces to showing that there exists $b_0 > 0$ large enough such that, for every $0 < a \leq b$ with $b \geq b_0$, there exists $(\alpha, d) \in (0,1) \times \mathbb{R}_+^*$ and $\phi_0 \in (0, \pi)$ such that $a = \nu K_+(\phi_0)$, $b = \nu K_-(\phi_0)$ and (90) holds true. First, $\phi_0 \in (0, \pi)$ is determined as the largest solution in $(0, \pi)$ of (102), which always exists since the range of the function K_+/K_- is equal to $\mathbb{R}_+^*$. Since $a/b \leq 1$ and taking the bounds on K_-, one gets that $K_+(\phi_0)$ is bounded above by a universal constant and hence every solution of (102) is bounded below by a positive constant independent of $a \leq b$. Finally, since K_+ is stricly increasing, establishing (103) simply amounts to verify that $c_{\phi_0/2}^2 b/a$ admits a positive limit as a/b tends to zero. In the latter case, one necessarily has that $K_+(\phi_0)$ must tend to zero and then, by taking into account (107), the solution ϕ_0 is unique and tends to π so that

$$c_{\phi_0/2}^2 \sim_{a/b \to 0} \left(\frac{\pi - \phi_0}{2} \right)^2 \sim_{a/b \to 0} \frac{\sqrt{2}}{\pi} K_+(\phi_0) \sim_{a/b \to 0} \frac{a}{\sqrt{2}b}.$$

This proves (103).

We now show the existence of (α, d). Define $(A, D) \in (0,1) \times \mathbb{R}_+^*$ by

$$A = \frac{K_-(\phi_0)^2}{2b^2}, \quad \text{and} \quad D = \frac{(1 + c_{\phi_0}) K_-(\phi_0)^2}{4b^2}.$$

Then, by definition of ϕ_0, one necessarily has that

$$A = \frac{\alpha(1-\alpha)}{(1-\alpha+d)^2}, \quad \text{and} \quad D = \frac{d(1+d)}{(1-\alpha+d)^2}. \tag{113}$$

The previous equations can be written as $(A, D) = G(\alpha, d)$. Since the Jacobian matrix of G at $(0,0)$ is equal to Id_2, one can apply the inverse function theorem in a neighborhood $\mathcal{O}$ of $(0,0)$. Then, b_0 is chosen large enough so that (A, D) defined in (113) belongs to $\mathcal{O}$ for $b \geq b_0$ and $\phi_0 \in (0, \pi)$. It is then immediate to get (104) from (113) since both α and d tend to zero as b tends to infinity. $\square$

5.3. Proof of Proposition 21

Taking into account Proposition 6, it is enough to establish Proposition 21 for sequences $(a_l, b_l)_{l \in \mathbb{N}}$ such that b_l tends to infinity as l tends to infinity. Moreover, since we need to upper bound $\mu(a,b,2)$, it is enough to find a control $c \in \mathcal{C}_2^0(a,b,a+b)$ and an initial condition $\omega_0 \in \mathbb{S}^1$, whose cost $J(c,\omega_0)$ is indeed smaller than $C_0 a/(1+b^2)$ for some universal constant C_0. We claim that such a control is provided by Proposition 24. Showing this claim simply amounts to compute the cost of such a control and to verify the desired inequality, as performed in the next result.

Lemma 28. *Assume that b is large enough. Let $\lambda = (\theta, \mathrm{diag}(a,b), \eta, \alpha, d)$ be the extremal satisfying Assumption 1 and associated with the control $S = cc^\top$ with $c = e^{i\phi/2} \in \mathcal{C}_2^0(a,b,a+b)$ and initial condition ω_0 provided by Proposition 24. Then, there exist two positive constants c_2, c_3 independent of (a,b) for b large enough such that*

$$c_2 \frac{a}{b^2} \le J(c,\omega_0) \le c_3 \frac{a}{b^2}. \tag{114}$$

Proof. From (OCP), one has at once that

$$J(c,\omega_0) = \int_0^{a+b} c^2_{(\theta-\phi)/2} dt, \tag{115}$$

and $c^2_{(\theta-\phi)/2} = s^2_{\varepsilon/2}$, where $\varepsilon = \theta - \psi - \pi$. From (Σ), the dynamics of ε on $[0,a+b]$ is given by

$$\ddot{\varepsilon} = -\mu s_\varepsilon + s_\varepsilon c_\varepsilon, \quad \text{where} \quad \mu = \frac{1}{1-\alpha+d}. \tag{116}$$

Moreover, the initial conditions $(\varepsilon_0, \dot{\varepsilon}_0) = (\varepsilon(0), \dot{\varepsilon}(0))$ are solutions of

$$c_\varepsilon = 1 - \frac{2\alpha d}{1-\alpha+d}, \qquad \dot{\varepsilon} = -s_\varepsilon = -(\alpha+d) s_{\phi_0}, \tag{117}$$

and $(\varepsilon(a+b), \dot{\varepsilon}(a+b)) = (-\varepsilon_0, -\dot{\varepsilon}_0)$.

First notice that (116) can be written $\ddot{\varepsilon} = -s_\varepsilon(\mu - c_\varepsilon)$, yielding that $\ddot{\varepsilon}$ and s_ε have opposite signs since $\mu > 1$ according to Remark 25. Moreover, if there exists $t_1 \in [0,a+b]$ such that $\varepsilon(t_1) = 0$, then $\varepsilon(t_1+t) = -\varepsilon(t_1-t)$ for times t_1-t, t_1+t in $[0,a+b]$.

We have the following first integral for ε after integrating between the times zero and $t \in [0,a+b]$ and taking into account (117):

$$\dot{\varepsilon}^2 = 2\mu(c_\varepsilon - c_{\varepsilon_0}) + s_\varepsilon^2. \tag{118}$$

Since ε starts at time $t=0$ with negative speed $\dot{\varepsilon}_0$ according to (117), ε will decrease in a right neighborhood of $t=0$. Note that, from (118), $\dot{\varepsilon}$ will keep the same sign, i.e., negative, as long as $|\varepsilon| \le \varepsilon_0$. Hence, ε will reach the value $\varepsilon = -\varepsilon_0$ at a time t_0, however with a negative speed. Therefore, by (117), t_0 must be strictly smaller than $a+b$ and ε decreases in a right neighborhood of $t=t_0$. This will go on till either $\dot{\varepsilon} = 0$ or $\varepsilon = -\pi/2$, since at time $t = a+b$ we have $\varepsilon(a+b) = -\varepsilon_0$ and $\dot{\varepsilon}(a+b) = -\dot{\varepsilon}_0$. The latter possibility is clearly ruled out since the r.h.s. of (118) is negative at $\varepsilon = -\pi/2$ for b sufficiently large. Then, $\dot{\varepsilon} = 0$ occurs at some time $\bar{t} < a+b$ for the $\varepsilon = -\bar{\varepsilon}$, where $\bar{\varepsilon}$ is the unique angle in $(0,\pi/2)$ verifying

$$2\mu(c_{\bar{\varepsilon}} - c_{\varepsilon_0}) + s^2_{\bar{\varepsilon}} = 0. \tag{119}$$

Since $\ddot{\varepsilon}(\bar{t}) > 0$ (because $s_{\varepsilon(\bar{t})} < 0$), then $\bar{t}$ is an isolated zero of $\dot{\varepsilon}$ and the latter must change sign there, implying that ε increases in a right neighborhood of $t = \bar{t}$. By a similar reasoning as before, ε increases till $\varepsilon = -\varepsilon_0$ at a time τ and one will also get that $\dot{\varepsilon}(\tau) = \dot{\varepsilon}_0 = -s_{\varepsilon(\tau)}$. Then either $\tau = a + b$ or ε increases in a right neighborhood of $t = \tau$. By repeating the above argument, one can see that ε will oscillate between $-\bar{\varepsilon}$ and $\bar{\varepsilon}$ and stop when (117) is satisfied. In any case, it must hold that $b + a = m\tau$ for some positive integer m. We show next that $m = 1$.

For that purpose, let us first prove the following asymptotics for $\bar{\varepsilon}$

$$\bar{\varepsilon} \sim 2\sqrt{d}, \qquad \text{as } b \to +\infty. \tag{120}$$

We can rewrite (119) as

$$(\mu - c_{\bar{\varepsilon}})^2 = 1 + \mu^2 - 2\mu c_{\varepsilon_0}.$$

Then, by the explicit expression of μ and the fact that $\alpha > d$, we have

$$c_{\bar{\varepsilon}} = \mu - \left(1 + \mu^2 - 2\mu c_{\varepsilon_0}\right)^{1/2} = 1 - \frac{2d}{1 - \alpha + d}, \tag{121}$$

which implies (120).

By the previous claim, $|\varepsilon| \leq 3\sqrt{d}$ for b large enough on $[0, a+b]$. Subtracting (119) to (118) yields

$$\dot{\varepsilon}^2 = 2\mu(c_\varepsilon - c_{\bar{\varepsilon}}) - c_\varepsilon^2 + c_{\bar{\varepsilon}}^2 = (c_\varepsilon - c_{\bar{\varepsilon}})\big((\mu - c_\varepsilon) + (\mu - c_\varepsilon)\big).$$

We have $\mu - c_\varepsilon = (\alpha + d)\big(1 + O(\alpha)\big)$ and $\mu - c_\varepsilon = \alpha - d + 2s_{\varepsilon/2}^2$. Hence

$$(\mu - c_\varepsilon) + (\mu - c_\varepsilon) = (2\alpha + \varepsilon^2/2)\big(1 + O(\alpha)\big).$$

On the other hand,

$$c_\varepsilon - c_{\bar{\varepsilon}} = \frac{\bar{\varepsilon}^2 - \varepsilon^2}{2}\big(1 + O(\alpha)\big).$$

Gathering the previous inequalities then yields

$$\dot{\varepsilon}^2 = (\bar{\varepsilon}^2 - \varepsilon^2)(\alpha + (\varepsilon/2)^2)(1 + O(\alpha)), \tag{122}$$

where $O(\alpha)$ denotes a function of the time such that $|O(\alpha)| \leq c\alpha$ on the interval $[0, T]$, for some $c > 0$ independent of b.

We now prove that $\tau = a + b$. By obvious symmetry considerations of solutions of (116), one has that

$$\tau = \int_0^\tau dt = 2\int_0^{\bar{\varepsilon}} \frac{d\varepsilon}{\big(2\mu(c_\varepsilon - c_{\bar{\varepsilon}}) - c_\varepsilon^2 + c_{\bar{\varepsilon}}^2\big)^{1/2}}.$$

It is then easy to see that τ is a continuous function of (a, b), for b large enough. On the other hand, by construction, the continuous function $(a, b) \mapsto m = (a + b)/\tau$ takes only positive integer values, and hence m must be independent of (a, b) for b large enough. In particular, it holds

$$m = \lim_{b\to+\infty} \frac{a+b}{\tau}, \qquad \text{for every } a > 0. \tag{123}$$

In order to compute the above limit, we observe that, using (120), we have

$$\tau \sim 2\big(1 + O(\alpha)\big) \int_0^{\bar\varepsilon} \frac{d\varepsilon}{\big((\bar\varepsilon^2 - \varepsilon^2)(\alpha + \frac{\varepsilon^2}{2})\big)^{1/2}}, \qquad \text{as } b \to +\infty.$$

Thus, again by (120), one gets that $\tau \sim \big(1 + O(\alpha)\big) I(a, b)$, where

$$I(a, b) = 2 \int_0^1 \frac{dv}{\big((1 - v^2)(\alpha + dv^2)\big)^{1/2}}.$$

Since $a > 0$ is fixed, letting b tend to infinity we have $a/b \to 0$. Hence,

$$I(a, b) \sim \frac{2}{\sqrt{\alpha}} \int_0^1 \frac{dv}{(1 - v^2)^{1/2}} = \frac{\pi}{\sqrt{\alpha}} \sim \frac{1}{\sqrt{\alpha}} \frac{\sqrt{2}\pi b}{K_-(\phi_0)} \sim b, \qquad \text{as } b \to +\infty$$

By (123) this yields that $m = 1$ and hence that $\tau = a + b$ for every b sufficiently large, independently of a.

We can finally prove the desired estimate for the cost. Indeed, by taking into account the previous results, one has

$$J(c, \omega_0) = 2 \int_0^{\bar\varepsilon} \frac{s^2_{\varepsilon/2}\, d\varepsilon}{\big(2\mu(c_\varepsilon - c_{\bar\varepsilon}) - c^2_{\varepsilon} + c^2_{\bar\varepsilon}\big)^{1/2}}.$$

By the same arguments as before, this yields

$$J(c, \omega_0) \sim \frac{(1 + O(\alpha))}{2} \int_0^{\bar\varepsilon} \frac{\varepsilon^2 d\varepsilon}{\big((\bar\varepsilon^2 - \varepsilon^2)(\alpha + \frac{\varepsilon^2}{2})\big)^{1/2}}, \qquad \text{as } b \to +\infty.$$

Thanks to (120), this further simplifies to $J(c, \omega_0) \sim (1 + O(\alpha)) \frac{d}{\sqrt{\alpha}} \mathcal{J}(d/\alpha)$, where

$$\mathcal{J}(\gamma) = 2 \int_0^1 \frac{v^2 dv}{\big((1 - v^2)(1 + 2\gamma v^2)\big)^{1/2}}, \qquad \gamma > 0.$$

Since $d/\alpha \in (0, 1)$, $\mathcal{J}(d/\alpha)$ is bounded below and above by positive constants independent of (a, b). Together with Remark 25, this concludes the proof of the statement. $\square$

A. Hamiltonian equations

In this appendix, we apply the Pontryagin Maximum Principle (PMP) to the control system (Conv), in order to derive necessary optimality conditions. These are essential to the proofs of Proposition 16.

For the Hamiltonian formalism used below, we refer to [1].

Lemma 29. *Let $H \in C^\infty(T^*\mathbb{S}^{n-1})$ be an Hamiltonian function. Upon the identification $T^*_\omega\mathbb{S}^{n-1} \simeq \omega^\perp$, the corresponding Hamiltonian system $\dot\xi = \vec{H}(\xi)$, $\xi = (\omega, p) \in T^*\mathbb{S}^{n-1}$, reads*

$$\dot\omega = \frac{\partial H}{\partial p} - \left(\omega^\top \frac{\partial H}{\partial p}\right)\omega \tag{124}$$

$$\dot p = \frac{\partial H}{\partial \omega} - \left(\omega^\top \frac{\partial H}{\partial \omega}\right)\omega - \left(\omega^\top \frac{\partial H}{\partial p}\right)p + \left(p^\top \frac{\partial H}{\partial p}\right)\omega. \tag{125}$$

Proof. Upon the given identifications, we have that

$$T_{(\omega,p)}(T^*\mathbb{S}^{n-1}) = \left\{(v_1, v_2) \in \mathbb{R}^{2n} \mid \omega^\top v_1 = 0 \quad \text{and} \quad p^\top v_1 + \omega^\top v_2 = 0\right\}. \tag{126}$$

Letting $(\frac{\partial H}{\partial \omega}, \frac{\partial H}{\partial p}) \in \mathbb{R}^{2n}$ be the partial derivative at $\xi = (\omega, p) \in \mathbb{S}^{n-1}$ of H, we have

$$d_\xi H(v_1, v_2) = v_1^\top \frac{\partial H}{\partial \omega} + v_2^\top \frac{\partial H}{\partial p}, \qquad \forall (v_1, v_2) \in T_{(\omega,p)}(T^*\mathbb{S}^{n-1}). \tag{127}$$

On the other hand, the Hamiltonian vector field $\vec{H} \in \Gamma(T^*\mathbb{S}^{n-1})$, with components $\vec{H} = (\vec{H}_p, -\vec{H}_\omega) \in T(T^*\mathbb{S}^{n-1})$, is the only vector field such that

$$d_\xi H(v_1, v_2) = v_1^\top \vec{H}_\omega(\xi) + v_2^\top \vec{H}_p(\xi), \qquad \forall (v_1, v_2) \in T_{(\omega,p)}(T^*\mathbb{S}^{n-1}). \tag{128}$$

As a consequence of these two facts, we have

$$\omega^\top \vec{H}_p(\omega, p) = 0, \qquad p^\top \vec{H}_p(\omega, p) = \omega^\top \vec{H}_\omega(\omega, p), \tag{129}$$

$$v_1^\top \left(\frac{\partial H}{\partial \omega} - \vec{H}_\omega(\omega, p)\right) + v_2^\top \left(\frac{\partial H}{\partial p} - \vec{H}_p(\omega, p)\right) = 0, \qquad \forall (v_1, v_2) \in T_{(\omega,p)}(T^*\mathbb{S}^{n-1}). \tag{130}$$

By (126), we have that, if $(v_1, v_2) \in T_{(\omega,p)}(T^*\mathbb{S}^{n-1})$ is such that $\omega^\top v_2 = 0$, then $v_1 = 0$. As a consequence, considering (130) for such (v_1, v_2) and taking into account the first equation of (129), yields

$$\vec{H}_p(\omega, p) = \frac{\partial H}{\partial p} - \left(\omega^\top \frac{\partial H}{\partial p}\right)\omega. \tag{131}$$

Plugging this in (130), one deduces that

$$\vec{H}_\omega(\omega, p) = \frac{\partial H}{\partial \omega} - \left(\omega^\top \frac{\partial H}{\partial \omega}\right)\omega - \left(\omega^\top \frac{\partial H}{\partial p}\right)p + \left(p^\top \frac{\partial H}{\partial p}\right)\omega. \tag{132}$$

This completes the proof. ∎

Proposition 30. *Let* $(\omega, Q) : [0,T] \to \mathcal{M}$ *be an optimal trajectory of system* (Conv)*, associated with an optimal control* S*. Then, there exists a curve* $t \in [0,T] \mapsto (p(t), P_Q(t)) \in T^*_{\omega(t)}\mathbb{S}^{n-1} \times T^*_{Q(t)} \operatorname{Sym}_n$ *and* $\nu_0 \in \{0,1\}$*, satisfying* (42)*,* (43)*,* (44)*, and* (45)*.*

Proof. Let $\lambda = (\omega, Q, p, P_Q)$. Recall that we consider the identification $T^*_\omega \mathbb{S}^{n-1} \times T^*_Q \operatorname{Sym}_n \simeq (\mathbb{R}\omega)^\perp \times \operatorname{Sym}_n$. We also let $T_{(\omega,Q)}\mathcal{M} \simeq (\mathbb{R}\omega)^\perp \times \operatorname{Sym}_n$, so that

$$\langle \lambda, v\rangle = p^\top v + \operatorname{Tr}(P_Q V) = \operatorname{Tr}(vp^\top + VP_Q), \qquad \text{for every } v = (v, V) \in T_{(\omega,Q)}\mathcal{M}. \tag{133}$$

We follow the formulation of the PMP given in [1, Theorem 12.4]. Simple computations show that, for $\nu_0 \in \{0,1\}$, the Hamiltonian associated with the system is given by (up to constants)

$$H^{\nu_0}(\lambda, S) = \frac{\operatorname{Tr}(S\tilde{M})}{2}, \quad \text{where} \quad \tilde{M} = P_Q - \left(\omega p^\top + p\omega^\top + \nu_0 \omega\omega^\top\right). \tag{134}$$

Equations (44) and (45) are immediate consequences of the PMP. In order to complete the proof, we are hence left to check (42) and (43). By the PMP, we have $\dot{\lambda} = \vec{H}$, where we let $\vec{H} \in \Gamma(T^*\mathcal{M})$ be the Hamiltonian vector field associated with the control S.

We observe that the Hamiltonian decomposes as $H(\omega, Q, p, P_Q, S) = H_1(\omega, p, S) + H_2(Q, P_Q, S)$. This implies that a similar decomposition holds for the corresponding Hamiltonian vector field. Thus, (42) follows by Lemma 29 and the fact that

$$\frac{\partial H_1}{\partial \omega} = -Sp - \nu_0 S\omega, \quad \text{and} \quad \frac{\partial H_1}{\partial p} = -S\omega. \tag{135}$$

On the other hand, (43) follows from the easily verified fact that

$$\vec{H}_2 = \left(\frac{\partial H_2}{\partial P_Q}, -\frac{\partial H_2}{\partial Q}\right) = (S, 0).$$

This concludes the proof of the statement. □

Proposition 31. *Let* $\mathcal{M} = \mathbb{S}^{n-1} \times \operatorname{Sym}_n \times O(n)$*, and consider the optimal control problem* (P-Conv) *defined in the proof of Proposition 17. Then, every extremal of this problem has to satisfy* (52)*.*

Proof. A decomposition argument similar to the one used in the preceding proof, allows one to reduce to the case where $\mathcal{M} = O(n)$ and $H(U, p_U, S) = \operatorname{Tr}(USU^\top P_Q)$ for some $P_Q \in Sy_n^+$. Then, for every $(U, p_U) \in T\mathcal{M}$, the tangent space $T_{(U,p_U)}(T\mathcal{M})$ can be identified with

$$\left\{(A,B) \in M_n(\mathbb{R}) \times M_n(\mathbb{R}) \mid \left(U^\top A, U^\top B - \frac{U^\top p_U U^\top A + U^\top A U^\top p_U}{2}\right) \in \operatorname{Skew}_n^2\right\}. \tag{136}$$

Simple computations yield

$$\frac{\partial H}{\partial U} = P_Q US, \quad \text{and} \quad \frac{\partial H}{\partial p_U} = 0. \tag{137}$$

One therefore deduces that $\vec{H} = (0, U[S, U^\top P_Q U])$, completing the proof. □

References

[1] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[2] B. Anderson. Exponential stability of linear equations arising in adaptive identification. *IEEE Transactions on Automatic Control*, 22(1):83–88, February 1977.

[3] B. Anderson, R. R. Bitmead, C. R. Johnson Jr, P. V. Kokotovic, R. L. Kosut, I. M. Mareels, L. Praly, and B. D. Riedle. *Stability of adaptive systems: Passivity and averaging analysis*. MIT press, 1986.

[4] S. Andersson and P. S. Krishnaprasad. Degenerate gradient flows: a comparison study of convergence rate estimates. In *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, volume 4, pages 4712–4717 vol.4, Dec 2002.

[5] V. I. Arnol' d. *Ordinary differential equations*. MIT Press, Cambridge, Mass.-London, 1978. Translated from the Russian and edited by Richard A. Silverman.

[6] N. Barabanov and R. Ortega. On global asymptotic stability of $\dot{x} = -\phi(t)\phi^\top(t)x$ with ϕ not persistently exciting. *Systems and Control Letters*, 109:24–29, 2017.

[7] N. Barabanov, R. Ortega, and A. Astolfi. Is normalization necessary for stable model reference adaptive control? *IEEE Transactions on Automatic Control*, 50(9):1384–1390, Sep. 2005.

[8] R. Brockett. The rate of descent for degenerate gradient flows. In *Proceedings of the 2000 MTNS, Perpignan, France.*, Jun 2000.

[9] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[10] D. Efimov, N. Barabanov, and R. Ortega. Robust stability under relaxed persistent excitation conditions. In *CDC 2018 - 57th IEEE Conference on Decision and Control*, Fontainebleau (FL), United States, Dec. 2018.

[11] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[12] L. Praly. Convergence of the gradient algorithm for linear regression models in the continuous and discrete time cases. Research report, PSL Research University ; Mines ParisTech, Feb. 2017.

[13] A. Rantzer. To estimate the L_2-gain of two dynamic systems. In *Open problems in mathematical systems and control theory*, Comm. Control Engrg. Ser., pages 177–179. Springer, London, 1999.

[14] M. M. Sondhi and D. Mitra. New results on the performance of a well-known class of adaptive filters. *Proceedings of the IEEE*, 64(11):1583–1597, Nov 1976.

[15] A. Weiss and D. Mitra. Digital adaptive filters: Conditions for convergence, rates of convergence, effects of noise and errors arising from the implementation. *IEEE Transactions on Information Theory*, 25(6):637–652, November 1979.