ÉCOLE POLYTECHNIQUE
PROMOTION X2008

Guilherme MAZANTI

Rapport de stage de recherche

Stabilisation des systèmes linéaires à excitation persistante

Non confidentiel

Option : Mathématiques Appliquées
Champ de l'option : MAP593 - Automatique et Recherche Opérationnelle
Directeur de l'option : Yacine CHITOUR
Directeurs de stage : Yacine CHITOUR et Mario SIGALOTTI
Dates du stage : 06/04/2011 à 24/06/2011
Nom et adresse de l'organisme : École Polytechnique
CMAP - Centre de Mathématiques Appliquées
Route de Saclay
91128 PALAISEAU CEDEX, France

Abstract

We study the control system $\dot{x} = Ax + \alpha(t)bu$ where the pair (A,b) is controllable, $x \in \mathbb{R}^d$, $u \in \mathbb{R}$ is a scalar control and the unknown signal $\alpha : \mathbb{R}_+ \to [0,1]$ is persistently exciting (PE), i.e., there exists $T \geq \mu > 0$ such that, for all $t \in \mathbb{R}_+$, $\int_t^{t+T} \alpha(s)ds \geq \mu$. We are interested in the problem of stabilization at an arbitrary rate of convergence of this system by a linear state feedback $u = -Kx$. We start by recalling the main results already obtained for this kind of system, and that stabilization at an arbitrary rate is not possible for general PE signals when A is the double integrator. We thus restrict the class of PE signals considered by taking only M-lipschitzian PE signals α. In this case, we can show that the double integrator can be stabilized at an arbitrary rate of convergence.

Résumé

On étudie le système de commande $\dot{x} = Ax + \alpha(t)bu$ où la paire (A,b) est commandable, $x \in \mathbb{R}^d$, $u \in \mathbb{R}$ est une commande scalaire et le signal inconnu $\alpha : \mathbb{R}_+ \to [0,1]$ est à excitation persistante (PE), i.e., il existe $T \geq \mu > 0$ tels que, pour tout $t \in \mathbb{R}_+$, $\int_t^{t+T} \alpha(s)ds \geq \mu$. On s'intéresse au problème de stabilisation à taux de convergence arbitraire de ce système par un retour d'état linéaire $u = -Kx$. Nous commençons par rappeler les résultats principaux déjà obtenus pour ce type de système, notamment que la stabilisation à taux de convergence arbitraire n'est pas possible pour des signaux PE généraux lorsque A est le double intégrateur. Nous restreignons ainsi la classe de signaux PE étudiée en ne considérant que les signaux PE M-lipschitziens α. Dans ce cas, on peut montrer que le double intégrateur peut être stabilisé à un taux de convergence arbitraire.

Resumo

Estudamos o sistema de controle $\dot{x} = Ax + \alpha(t)bu$ em que o par (A,b) é controlável, $x \in \mathbb{R}^d$, $u \in \mathbb{R}$ é um controle escalar e o sinal desconhecido $\alpha : \mathbb{R}_+ \to [0,1]$ é um sinal a excitação persistente (PE), i.e., existem $T \geq \mu > 0$ tais que, para todo $t \in \mathbb{R}_+$, $\int_t^{t+T} \alpha(s)ds \geq \mu$. Interessamo-nos ao problema de estabilização a taxa de convergência arbitrária deste sistema através de uma realimentação de estado linear $u = -Kx$. Começamos retomando os principais resultados já obtidos para este tipo de sistema, em particular que a estabilização a velocidade arbitrária não é possível para sinais PE genéricos no caso em que A é o duplo integrador. Restringimos assim a classe de sinais PE estudada, considerando apenas os sinais PE M-lipschitzianos α. Neste caso, pode-se mostrar que o duplo integrador pode ser estabilizado a uma taxa de convergência arbitrária.

Contents

1 Introduction

The aim of the project was to continue the study of persistently excited (PE) linear control systems of [3, 4]. These systems can be written in the form

$$\dot{x} = Ax + \alpha(t)Bu$$

where $x \in \mathbb{R}^d$, $u \in \mathbb{R}^m$ is the control and $\alpha : \mathbb{R}_+ \to [0,1]$ is a scalar persistently exciting signal, that is, there exists two constants T, μ with $T \geq \mu > 0$ such that, for every $t \in \mathbb{R}_+$,

$$\int_t^{t+T} \alpha(s)ds \geq \mu.$$

Throughout this document, we shall consider only the case where the control u is scalar, and thus the matrix B is actually a column vector $b \in \mathbb{R}^d$.

This project was developed as the research project in the end of the third year of academic studies at École Polytechnique.

1.1 Non-autonomous linear systems

The study of linear systems of differential equations has a great interest, from both applied and theoretical points of view. These systems arise from linear models, as well as linearizations of non-linear systems at equilibrium points or along solutions. Consider a system in the form

$$\dot{x} = A(t)x \tag{1.1}$$

where $x \in \mathbb{R}^d$ and $t \mapsto A(t)$ is a bounded measurable application from a time interval I to the set of square $d \times d$ matrices with real coefficients, $\mathcal{M}_d(\mathbb{R})$. When A is constant, the system is called autonomous, and non-autonomous otherwise. By Carathéodory Theorem, given an initial condition $x(t_0) = x_0$ for a certain $t_0 \in I$, System (1.1) has a unique absolutely continuous solution, defined over the whole interval I. We shall be interested in the case $I = \mathbb{R}_+$.

The most simple case of non-autonomous linear system is when $A(t)$ may take only two values, A_1 and A_2, and another simple case is when $A(t)$ takes its values on the convex envelope of the set $\{A_1, A_2\}$, i.e., $A(t) \in [A_1, A_2] = \{(1-\alpha)A_1 + \alpha A_2, \alpha \in [0,1]\}$, A_1 and A_2 being two fixed matrices of $\mathcal{M}_d(\mathbb{R})$. The system may hence be written in the form

$$\dot{x} = ((1-\alpha(t))A_1 + \alpha(t)A_2)x \tag{1.2}$$

with $\alpha(t) \in \{0,1\}$ or $\alpha(t) \in [0,1]$. Such a system is thus able to switch between the dynamics of A_1 and A_2.

1.2 Control systems

The theory of linear autonomous control systems is concerned with the study of systems of the form

$$\dot{x} = Ax + Bu$$

with $x \in \mathbb{R}^d$, $u \in \mathbb{R}^m$, $A \in \mathcal{M}_d(\mathbb{R})$ and $B \in \mathcal{M}_{d,m}(\mathbb{R})$. The function u is the control of the system, which we choose trying to make the system have a certain prescribed behavior. A usual way to choose this control is on the form of a linear state feedback $u = -Kx$ with $K \in \mathcal{M}_{m,d}(\mathbb{R})$. This choice corresponds to replacing the original dynamics of the non-controlled system, given by the matrix A, with a new dynamics, given by the matrix $A - BK$.

1.3 Switched control systems

We can now combine the idea of replacing the dynamics of a system given by a certain matrix A with a new controlled dynamics $A - BK$ shown in Section 1.2 and the idea of a linear system switching between two dynamics as in (1.2): by taking $A_1 = A$ and $A_2 = A - BK$ on (1.2), we get

$$\dot{x} = (A - \alpha(t)BK)x$$

with $\alpha(t) \in [0,1]$. More generally, we are interested in the switched control system

$$\dot{x} = Ax + \alpha(t)Bu. \tag{1.3}$$

It is thus a system where the control u cannot act at every time t, but only when $\alpha(t)$ is active. System (1.3) is a linear switched control system and is the only case of switched control system that we shall study in this document; for a more complete reference, see, for instance, [6].

Physically, $\alpha(t)$ usually models an uncertainty on when the control can actually act on the system. Let us consider the example of a pendulum subject to a gravitational field g and excited by an external force u that can act only when the angle θ of pendulum is between two values θ_1 and θ_2, as shown on Figure 1.1. In this case, if we choose unitary mass, length of the pendulum and gravitational field, the corresponding control system is

$$\frac{d}{dt}\begin{pmatrix}\theta \\ \dot{\theta}\end{pmatrix} = \begin{pmatrix}\dot{\theta} \\ -\sin\theta\end{pmatrix} + \alpha(\theta)\begin{pmatrix}0 \\ -u\end{pmatrix}, \tag{1.4}$$

with $\alpha(\theta) = 1$ if $\theta_1 \leq \theta \leq \theta_2$ and $\alpha(\theta) = 0$ otherwise.

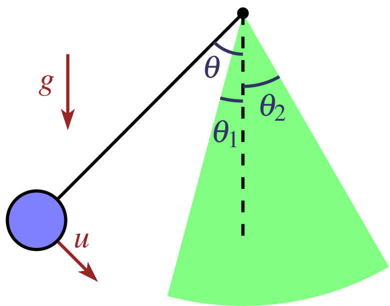

Figure 1.1: A simple pendulum excited by a force u. We suppose that the force can act only when the pendulum is on the region $\theta_1 \leq \theta \leq \theta_2$.

System (1.4) is a non-linear control system, and its non-linearities make it difficult to find an explicit solution of the system. Furthermore, we may be interested in situations where the angles θ_1 and θ_2 are not precisely known, where we know only estimates of these angles or where these angles may change with time or some other parameters. In these cases, it is interesting to look for estimations and properties of System (1.4) valid for a whole class $\mathcal{G}$ of functions α, and not only for a specific fixed function α. We suppose that α models a time-dependent uncertainty of when the control acts on the system, and we thus consider α to be a time-dependent function. Hence System (1.4) can be written as

$$\frac{d}{dt}\begin{pmatrix}\theta \\ \dot{\theta}\end{pmatrix} = \begin{pmatrix}\dot{\theta} \\ -\sin\theta\end{pmatrix} + \alpha(t)\begin{pmatrix}0 \\ -u\end{pmatrix} \tag{1.5}$$

with $\alpha \in \mathcal{G}$, and System (1.5) along with the hypothesis $\alpha \in \mathcal{G}$ is thus a switched control system.

1.4 Persistence of excitation

In the switched control system (1.3), the signal α determines when the control signal u may act. We suppose that α is not precisely known, and the only information on α we have is that it is chosen in a certain class of functions $\mathcal{G}$. The question that arises is which kind of class of functions $\mathcal{G}$ may be natural and useful in practice. It is obviously not interesting to choose $\mathcal{G} = L^\infty(\mathbb{R}_+, [0,1])$ since this class contains the constant function equal to 0 and functions which are arbitrarily small in norm L^∞ or that are different from 0 only in a finite time interval, and, in these cases, the control has a small or even null influence on the behavior of the non-controlled system $\dot{x} = Ax$. We are thus looking for a class of functions that are "active" often enough and that, at each activation, have a certain "minimum degree of activation". For this purpose, [3, 4] use signals α satisfying a condition called *persistent excitation* (PE): given two positive constants T, μ with $T \geq \mu$, we say that a measurable function $\alpha : \mathbb{R}_+ \to [0,1]$ is in the class $\mathcal{G}(T,\mu)$ if, for every $t \geq 0$,

$$\int_t^{t+T} \alpha(s) ds \geq \mu. \tag{1.6}$$

This condition is known as the condition of persistent excitation and appears naturally in the context of identification and adaptive control, as in [8]. Roughly speaking, it means that, for every time interval of length T, the amount of time that α is "active" during this time window is lower bounded by a positive constant.

1.5 Purpose of the project

The research project aimed at answering an open question raised in [4], its Open Problem 5. We consider the problem of stabilization to the origin, by means of a linear state feedback $u = -Kx$, of the persistently excited control system

$$\dot{x} = Ax + \alpha(t) bu \tag{1.7}$$

where $x \in \mathbb{R}^d$, $u \in \mathbb{R}$ and $\alpha \in \mathcal{G}(T,\mu)$. We want this stabilization to be uniform with respect to the signal α: K may depend on T and μ, but it should not depend on a particular signal $\alpha \in \mathcal{G}(T,\mu)$.

One may think that, intuitively, the stabilization of a system as (1.7) is possible under reasonable hypothesis on (A,b), due to the condition (1.6) imposed on α. Indeed, if we suppose that the linear control system defined by (A,b) is stabilizable, meaning that there exists $K^{\mathrm{T}} \in \mathbb{R}^d$ such that $A - bK$ is Hurwitz, and if we suppose, in order to simplify the reasoning, that $\alpha(t) \in \{0,1\}$, then, by choosing such a K, we can "stabilize" System (1.7) at least when $\alpha(t) = 1$, which, due to the condition (1.6) on α, happens at least for a total time μ on every time window of length T, but this is not actually a stabilization of the system since we still have to deal with time intervals when $\alpha(t) = 0$. We may then think that, if we know the behavior of the uncontrolled system $\dot{x} = Ax$, we can try to compensate any destabilization due to its dynamics when $\alpha(t) = 0$ by astutely choosing K so that the stabilization provided by $A - bK$ when $\alpha(t) = 1$ will eventually stabilize the system. This intuition can also be used in the case $\alpha(t) \in [0,1]$: one can prove (see [4]) that it is possible, for every $\rho > 0$, to choose a K that stabilizes System (1.7) for every $\alpha \in L^\infty(\mathbb{R}_+, [\rho,1])$. Then, for System (1.7) with a general $\alpha \in \mathcal{G}(T,\mu)$, by choosing ρ small enough, Equation (1.6) shows that we have $\alpha(t) \geq \rho$ at least for a total time that is uniformly lower bounded on every time window of length T; we have then "good" time intervals where α is lower bounded by ρ and we know how to stabilize the system, and "bad" time intervals where α is smaller than ρ. We then expect that, by studying the behavior of the system in the "bad" time intervals, we can choose K in order to compensate an eventual explosive behavior of the solution in the "bad" time intervals by convergence in the "good" ones in order to stabilize the system.

This intuition was proved true in [4], where it is shown that stabilization to the origin of System (1.7) is possible if (A,b) is a controllable pair and every eigenvalue of A has non-positive real part (see the precise statement in Theorem 3.2 below).

Let us just say that, even if we mentioned above quite quickly the study of the behavior of the system in "bad" intervals, this is actually a difficult point. All the difficulty comes from the fact that we want to choose a feedback $u = -Kx$ with K that is independent of $\alpha \in \mathcal{G}(T,\mu)$, and thus it is the same K that acts on both "good" and "bad" intervals. The feedback gain K is chosen in such a way that the behavior on "good" intervals is known and controlled, but we now nothing a priori on the effects of a particular choice of K on "bad" intervals, and obtaining estimates in this case is nontrivial, since the usual estimates, such as Gronwall's Lemma, are usually not fine enough to provide a useful result.

We can now ask the question of stabilization at an arbitrary rate: we want to choose a feedback $u = -Kx$ such that every solution of $\dot{x} = (A - \alpha(t)bK)x$ converges to 0 exponentially at a rate which is faster than a prescribed constant λ. In the case of a linear control system $\dot{x} = Ax + Bu$, the question is then to study the eigenvalues of $A - BK$, and the answer is given by the Pole Shifting Theorem, which guarantees that, for every controllable pair of matrices $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$ and every unitary polynomial P of degree d with real coefficients, there exists $K \in \mathcal{M}_{m,d}(\mathbb{R})$ such that P is the characteristic polynomial of $A - BK$, and then we can arbitrarily choose the eigenvalues of $A - BK$.

In the case $\alpha(t) \in \{0,1\}$, one might think that stabilization at an arbitrary rate still holds true for a controllable pair (A,b): it would suffice to choose K such that the solution of $\dot{x} = (A - bK)x$ stabilizes the system fast enough and that its convergence rate compensates any possible explosive behavior when $\alpha(t) = 0$. But this intuition is shown to be false in [4], where it is shown that, in dimension $d = 2$, there exists $\rho_\star$ such that, if $0 < \frac{\mu}{T} < \rho_\star$, then the maximal rate of convergence of System (1.7) is finite, which means that there exists a constant C such that, for every $K^{\mathrm{T}} \in \mathbb{R}^2$, there exists a signal $\alpha \in \mathcal{G}(T,\mu)$ and an initial condition x_0 of norm 1 such that the corresponding solution $x(t)$ of the system $\dot{x} = (A - \alpha bK)x$ satisfies

$$\limsup_{t \to +\infty} \frac{\log \|x(t)\|}{t} \geq -C. \tag{1.8}$$

The reason for not obtaining the intuitive idea is the overshoot phenomenon: when $\alpha(t) = 1$, one can actually choose K such that the solution of $\dot{x} = (A - bK)x$ stabilizes fast enough, but we recall that, by choosing such a K, even if the norm of the solutions of $\dot{x} = (A - bK)x$ tends to 0 with the desired convergence rate, it may increase a lot in a small interval of time $[0,t]$ before exponentially decreasing, a phenomenon that is known as overshoot. Then, if $\alpha(t) = 1$ for only a short period of time, it is actually the overshoot phenomenon, and not stabilization, that dominates the behavior of the solution of (1.7), which explains why the previous intuition is false. This is actually the idea of the proof of the previous result: for a given K, one constructs α such that the overshoot phenomenon dominates the stabilization provided by K.

Open Problem 5 raised in [4] is thus to recover stabilization by a linear state feedback at an arbitrary rate of convergence for System (1.7) by restricting the class $\mathcal{G}(T,\mu)$ where the signal α is taken. The proof of the impossibility of doing so for $\alpha \in \mathcal{G}(T,\mu)$ was done in [4] by explicitly constructing, for each $K^{\mathrm{T}} \in \mathbb{R}^2$, a signal α for which the solution $x(t)$ to the system $\dot{x} = (A - \alpha bK)x$ with a certain initial condition satisfies (1.8). The signal α constructed oscillates faster and faster as the norm of K increases. We may thus think that bounding the oscillations of α is a possibility to retrieve the arbitrary rate of convergence for System (1.7). More explicitly, we shall consider the subclass $\mathcal{D}(T,\mu,M)$ of $\mathcal{G}(T,\mu)$ of persistently exciting signals that are M-Lipschitz. The purpose of the project is thus to prove the stabilization at an arbitrary rate of convergence for System (1.7) when the function α is taken in $\mathcal{D}(T,\mu,M)$.

Intuitively, the Lipschitz hypothesis helps solving the problem created by the overshoot phenomenon that prevented stabilization at an arbitrary rate. In fact, if we choose K such that the solution of $\dot{x} = (A - bK)x$ stabilizes fast enough, we still have the overshoot phenomenon when $\alpha(t) = 1$ for a short interval of time, but now we cannot switch α to 0 right after the overshoot due to the Lipschitz hypothesis. Then, whenever α is larger than a certain positive constant ρ, it will stay larger than $\rho/2$ during an interval whose length is lower bounded, say, by ℓ, and thus we can hope to choose K in such a way that the stabilization provided by K will now dominate the destabilizing effect of the overshoot phenomenon whenever the length of the interval is greater than ℓ. This an the idea that motivates the choice of the class $\mathcal{D}(T,\mu,M)$.

1.6 Organization of the document

In Section 2, we shall present the notations and definitions used throughout this document. Section 3 then presents previous results on linear persistently excited control systems, especially those obtained in [3, 4], and including the precise statement of the results that we mentioned above.

We then turn in Section 4 to the core of the project, that is, the proof of a stabilization result at arbitrary rate on the case where α is taken in the class $\mathcal{D}(T,\mu,M)$. We prove that indeed, by restricting the class $\mathcal{G}(T,\mu)$ to $\mathcal{D}(T,\mu,M)$, we can retrieve the stabilization at an arbitrary rate, at least for the case of the double integrator. This proof is presented in several steps. We first do a change of variables that puts apart all the convergence information that we have for the system, and then the work we must do is to study the maximal divergence rate of the system in the new variables. To do so, we decompose the time $\mathbb{R}_+$ into two classes of time intervals, $\mathcal{I}_+$, the "good" intervals, where a certain function γ is larger than a certain positive number, and $\mathcal{I}_0$, the "bad" intervals, where γ is small, retrieving thus the idea of "good" and "bad" intervals mentioned above. Estimations on "good" intervals are done by integrating the equations of the movement in polar coordinates: if we take the feedback gain K large enough, we can show that the solution rotates around the origin in "good" intervals, and this fact can then be used in order to say that the norm of the solution is a function of the polar angle within a rotation, which is used to estimate the growth of the norm. A different strategy is necessary in "bad" intervals: we use the techniques of optimal control, and in particular Pontryagin Maximum Principle, in order to find the "worst trajectory", a particular solution of the system where we have the greatest growth rate possible for a "bad" interval, and we then use this particular solution to estimate the growth rate. The final part of the proof is to put all the results together and go back to the original variables in order to see that the convergence rate that we have from the change of variables is greater than the maximal explosion rate in the new variables if we have a large enough feedback gain K.

2 Notations and definitions

2.1 Basic definitions

In this document, $\mathcal{M}_{d,m}(\mathbb{R})$ denotes the set of $d \times m$ matrices with real coefficients. When $m = d$, this set is denoted simply by $\mathcal{M}_d(\mathbb{R})$. We identify the column matrices of $\mathcal{M}_{d,1}(\mathbb{R})$ with the vectors of $\mathbb{R}^d$ by the canonical identification. The euclidean norm of an element $x \in \mathbb{R}^d$ is denoted by $\|x\|$, and the associate matrix norm of a matrix $A \in \mathcal{M}_d(\mathbb{R})$ is also denoted by $\|A\|$, whereas the symbol $|x|$ is reserved for the absolute value of a real or complex number x. The real and imaginary parts of a complex number z are denoted respectively by $\Re(z)$ and $\Im(z)$.

2.2 PE and PEL signals and systems

We shall consider control systems of the form

$$\dot{x} = Ax + \alpha(t)Bu \tag{2.1}$$

where $x \in \mathbb{R}^d$, $A \in \mathcal{M}_d(\mathbb{R})$, $u \in \mathbb{R}^m$ is the control signal, $B \in \mathcal{M}_{d,m}(\mathbb{R})$ and α is a persistently exciting signal. We start by defining this notion.

Definition 2.1 (PE signal and (T,μ)-signal)**.** Let T, μ be two positive constants with $T \geq \mu$. We say that a measurable function $\alpha : \mathbb{R}_+ \to [0,1]$ is a (T,μ)-*signal* if, for every $t \in \mathbb{R}_+$, one has

$$\int_t^{t+T} \alpha(s)ds \geq \mu. \tag{2.2}$$

The set of (T,μ)-signals is denoted by $\mathcal{G}(T,\mu)$. We say that a measurable function $\alpha : \mathbb{R}_+ \to [0,1]$ is a *persistently exciting signal* (or simply *PE signal*) if it is a (T,μ)-signal for certain positive constants T and μ with $T \geq \mu$.

We shall use later on a restriction of this class, namely the Lipschitz (T,μ)-signals, which we define below.

Definition 2.2 (PEL signal and (T,μ,M)-signal)**.** Let T, μ and M be positive constants with $T \geq \mu$. We say that a measurable function $\alpha : \mathbb{R}_+ \to [0,1]$ is a (T,μ,M)-*signal* if it is a (T,μ)-signal and in addition α is globally M-Lipschitz, that is, for every $t,s \in \mathbb{R}_+$,

$$|\alpha(t) - \alpha(s)| \leq M\,|t-s|.$$

The set of (T,μ,M)-signals is denoted by $\mathcal{D}(T,\mu,M)$. We say that a measurable function $\alpha : \mathbb{R}_+ \to [0,1]$ is a *persistently exciting Lipschitz signal* (or simply *PEL signal*) if it is a (T,μ,M)-signal for certain positive constants T, μ and M with $T \geq \mu$.

We recall that a M-Lipschitz function α defined on $\mathbb{R}_+$ is in particular absolutely continuous, and hence it is differentiable almost everywhere, its derivative $\dot{\alpha}$ being bounded by M. Furthermore, α satisfies

$$\alpha(t) = \alpha(t') + \int_{t'}^{t} \dot{\alpha}(s)ds$$

for every $t,t' \in \mathbb{R}_+$; for more details, see, for instance, [9]. These results will be much used in what follows.

We can now define the object of our study, the persistently excited control system (2.1) given by a certain choice of matrices $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$ and constants T, μ and eventually the Lipschitz constant M.

Definition 2.3 (PE and PEL systems)**.** Given a pair $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$ and two positive constants T and μ (resp. three positive constants T, μ and M) with $T \geq \mu$, we say that the family of linear control systems

$$\dot{x} = Ax + \alpha Bu, \qquad \alpha \in \mathcal{G}(T,\mu) \quad (\text{resp. } \alpha \in \mathcal{D}(T,\mu,M)) \tag{2.3}$$

is the *PE system* associated with A, B, T and μ (resp. the *PEL system* associated with A, B, T, μ and M).

2.3 Controllability

An important property that we shall use is the controllability of a linear control system. We recall the general definition of controllability and Kalman's condition of controllability for a linear autonomous control system, which is a classic result on controllability and can be found in many reference books, such as [1, 10].

Definition 2.4 (Controllability)**.** We say that a control system given by the equation

$$\dot{x} = f(t,x,u), \qquad x \in \mathbb{R}^d,\, u \in \mathbb{R}^m \tag{2.4}$$

is *controllable in time* $T > 0$ if, for every $x_0, x_1 \in \mathbb{R}^d$, there exists a control signal $u : [0,T] \to \mathbb{R}^m$ such that the corresponding solution x of (2.4) is defined on $[0,T]$ and satisfies $x(0) = x_0$, $x(T) = x_1$. It is said to be *controllable* if it is controllable in time T for every $T > 0$.

Proposition 2.5 (Kalman's Criterion on Controllability)**.** *Consider the linear control system*

$$\dot{x} = Ax + Bu, \qquad x \in \mathbb{R}^d,\, u \in \mathbb{R}^m.$$

For every $T > 0$, this system is controllable in time $T > 0$ if and only if the controllability matrix $\mathcal{C}(A,B)$*, defined by*

$$\mathcal{C}(A,B) = \begin{pmatrix} B & AB & A^2B & \cdots & A^{d-1}B \end{pmatrix}$$

has rank d.

From now on, we shall suppose that the pair of matrices $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^d$ is controllable, which means that the linear control system $\dot{x} = Ax + Bu$ is controllable and hence that (A,B) satisfies Kalman's criterion.

2.4 Stabilizability

The main problem we are interested in is the question of stabilization of System (2.3) by a linear state feedback. A linear state feedback corresponds to the choice of the control u as $u = -Kx$ with $K \in \mathcal{M}_{m,d}(\mathbb{R})$, which makes System (2.3) take the form

$$\dot{x} = (A - \alpha(t)BK)x. \tag{2.5}$$

The problem of stabilization by a linear state feedback is thus the problem of choosing K such that the origin of the linear system (2.5) is globally asymptotically stable. With this in mind, we can define the notion of a stabilizer.

Definition 2.6 (Stabilizer)**.** Let T and μ (resp. T, μ and M) be positive constants with $T \geq \mu$. We say that $K \in \mathcal{M}_{m,d}(\mathbb{R})$ is a (T,μ)*-stabilizer* (resp. (T,μ,M)*-stabilizer*) for System (2.3) if, for every $\alpha \in \mathcal{G}(T,\mu)$ (resp. $\alpha \in \mathcal{D}(T,\mu,M)$), System (2.5) is globally asymptotically stable.

We remark that K may depend on T, μ and M, but it cannot depend on the particular signal $\alpha \in \mathcal{G}(T,\mu)$ or $\alpha \in \mathcal{D}(T,\mu,M)$. We also remark that a (T,μ)-stabilizer is also a (T,μ,M)-stabilizer for every $M > 0$.

The question we are interested in is not only to stabilize a PE or PEL system, but to stabilize it with an arbitrary rate of convergence. In order to rigorously define this idea, we introduce some concepts.

Definition 2.7. Let $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$, $K \in \mathcal{M}_{m,d}(\mathbb{R})$ and $T \geq \mu > 0$, $M > 0$, and consider System (2.5). Fix $\alpha \in \mathcal{G}(T,\mu)$ (resp. $\alpha \in \mathcal{D}(T,\mu,M)$). We denote by $x(t;x_0)$ the solution of System (2.5) with initial condition $x(0;x_0) = x_0$.

- The *maximal Lyapunov exponent* $\lambda^+(\alpha,K)$ associated with (2.5) is defined by

$$\lambda^+(\alpha,K) = \sup_{\|x_0\|=1} \limsup_{t\to+\infty} \frac{\log\|x(t;x_0)\|}{t}.$$

- The *rate of convergence* associated with the systems $\dot{x} = (A - \alpha(t)BK)x$, $\alpha \in \mathcal{G}(T,\mu)$ (resp. $\alpha \in \mathcal{D}(T,\mu,M)$) is defined as

$$\mathrm{rc}_{\mathcal{G}}(T,\mu,K) = -\sup_{\alpha\in\mathcal{G}(T,\mu)} \lambda^+(\alpha,K) \text{ (resp. } \mathrm{rc}_{\mathcal{D}}(T,\mu,M,K) = -\sup_{\alpha\in\mathcal{D}(T,\mu,M)} \lambda^+(\alpha,K)).$$

- The *maximal rate of convergence* associated with System (2.3) is defined as

$$\mathrm{RC}_{\mathcal{G}}(T,\mu) = \sup_{K\in\mathcal{M}_{m,d}(\mathbb{R})} \mathrm{rc}_{\mathcal{G}}(T,\mu,K) \text{ (resp. } \mathrm{RC}_{\mathcal{D}}(T,\mu,M) = \sup_{K\in\mathcal{M}_{m,d}(\mathbb{R})} \mathrm{rc}_{\mathcal{D}}(T,\mu,M,K)).$$

The stabilization of System (2.3) at an arbitrary rate of convergence corresponds thus to have $\mathrm{RC}_{\mathcal{G}}(T,\mu) = +\infty$ or $\mathrm{RC}_{\mathcal{D}}(T,\mu,M) = +\infty$.

The fact that we are interested in the maximal rate of convergence explains why we consider only the case where the pair $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$ is controllable, for, when it is not the case, the Pole Shifting Theorem shows that there are eigenvalues of $A-BK$ which do not depend on K, and then we do not have a result of stabilization at an arbitrary rate even in the case of a non-switched linear control system $\dot{x} = Ax+Bu$.

3 Previous results

3.1 Existence of a stabilizer

The first stabilization problem treated in [3] is the case of a neutrally stable system, that is, a system in the form (2.1) such that every eigenvalue of A has non-positive real part, and those with real part zero have trivial Jordan blocks. We consider the case where $(A,B) \in \mathcal{M}_d(\mathbb{R}) \times \mathcal{M}_{d,m}(\mathbb{R})$ is stabilizable, meaning that there exists $K \in \mathcal{M}_{m,d}(\mathbb{R})$ such that $A - BK$ is Hurwitz. In this case, the result is that the PE system

$$\dot{x} = Ax + \alpha(t)Bu, \qquad \alpha \in \mathcal{G}(T,\mu) \tag{3.1}$$

admits a (T,μ)-stabilizer K.

Theorem 3.1. *Assume that the pair (A,B) is stabilizable and that the matrix A is neutrally stable. Then there exists a matrix $K \in \mathcal{M}_{m,d}(\mathbb{R})$ such that, for every $T \geq \mu > 0$, K is a (T,μ)-stabilizer for* (3.1).

We remark that the theorem gives a gain K independent on T and μ. Also, since the result is true for the PE System (3.1), it is also true for a PEL system for every $M > 0$.

The proof of the theorem is given in details in [3]. We just recall here its main idea, which consists first to a reduction to the case where (A,B) is controllable and A is skew-symmetric, and then the proof of the result in this case is done by taking $K = B^{\mathrm{T}}$ and considering the Lyapunov function $V(x) = \frac{1}{2}\|x\|^2$. A technical lemma shows that, in this case, the Lyapunov function uniformly decreases in an interval $[t, t+T]$ for any $t > 0$, which gives the desired result.

The second case treated in [3] is the double integrator, which is generalized to the n-integrator and then to a more general case in [4]. The system considered is still (3.1), but we restrict ourselves to the case where the control u is scalar, which means that the matrix B is a column matrix $b \in \mathbb{R}^d$.

Theorem 3.2. *Let $(A,b) \in \mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^d$ be a controllable pair and assume that the eigenvalues of A have non-positive real part. Then, for every T, μ with $T \geq \mu > 0$, there exists a (T,μ)-stabilizer for* (3.1).

In the case of the double integrator, the system is given by the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and thus (3.1) is written as

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \alpha(t)u. \end{cases} \tag{3.2}$$

The proof in this case is based on the following fact: for every $\nu > 0$, $K = \begin{pmatrix} k_1 & k_2 \end{pmatrix}$ is a (T,μ)-stabilizer of (3.2) if and only if $\begin{pmatrix} \nu^2 k_1 & \nu k_2 \end{pmatrix}$ is a $(T/\nu, \mu/\nu)$-stabilizer of (3.2), which can be seen by considering the equation satisfied by

$$x_\nu(t) = \begin{pmatrix} 1 & 0 \\ 0 & \nu \end{pmatrix} x(\nu t).$$

The idea of the proof is thus to construct a $(T/\nu, \mu/\nu)$-stabilizer $K = \begin{pmatrix} k_1 & k_2 \end{pmatrix}$ for (3.2) for a certain ν large enough, and then the (T,μ)-stabilizer we seek for is $\begin{pmatrix} k_1/\nu^2 & k_2/\nu \end{pmatrix}$. The construction of such a K is based on a limit process: if we consider a family of signals $\alpha_n \in \mathcal{G}(T/\nu_n, \mu/\nu_n)$ with $\lim_{n\to+\infty} \nu_n = +\infty$, the weak-$\star$ compactness of $L^\infty(\mathbb{R}_+, [0,1])$ shows that this sequence admits a weak-$\star$ convergent subsequence in $L^\infty(\mathbb{R}_+, [0,1])$, converging weakly-$\star$ to a certain limit $\alpha_\star$, which can be shown to satisfy $\alpha_\star(t) \geq \frac{\mu}{T}$ almost everywhere. We can thus study the limit system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \alpha_\star(t)u, \end{cases} \qquad \alpha_\star(t) \geq \frac{\mu}{T}$$

in order to obtain properties of System (3.2) by a limit process. The general idea is thus to accelerate the dynamics of (3.2) by a factor ν. This acceleration reduces the importance of the intervals where α is small, since, in the limit, $\alpha_\star(t) \geq \frac{\mu}{T}$ almost everywhere, making it easier to study the behavior of the system. We thus construct a gain $K = \begin{pmatrix} k_1 & k_2 \end{pmatrix}$ for a system for which the accelerating scale ν is big enough, and then we are finally able to stabilize the original system by a small gain $\begin{pmatrix} k_1/\nu^2 & k_2/\nu \end{pmatrix}$.

We wrote the technique of accelerating the system and studying a limit system just for the case of the double integrator for simplicity, but these ideas can also be used in the general case. The difference is how to use the limit system to obtain a gain K for ν large enough: [3] uses many geometric properties of the system in $\mathbb{R}^2$, whereas [4] uses a more general technique based on Lyapunov functions. In both cases, however, the study of a limit system is essential to the proof. We also remark that these results, proved for a PE system, are also true for a PEL system as a particular case.

3.2 Stabilization at an arbitrary rate

As we remarked in Section 2.4, the problem of stabilization of a PE or PEL system at an arbitrary rate is formulated in terms of the maximal rates of convergence as the problem of determining whether $\mathrm{RC}_{\mathcal{G}}(T,\mu)$ and $\mathrm{RC}_{\mathcal{D}}(T,\mu,M)$ are finite or not. In this sense, [4] gives two results concerning the stabilization of PE systems.

Theorem 3.3. *Let d be a positive integer. There exists $\rho^\star \in (0,1)$ such that, for every controllable pair $(A,b) \in \mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^d$ and every positive T, μ satisfying $\rho^\star < \frac{\mu}{T} \leq 1$, one has $RC_{\mathcal{G}}(T,\mu) = +\infty$.*

This means that, at least for $\frac{\mu}{T}$ large enough, stabilization at an arbitrary rate of convergence is possible for a PE system with controllable (A,b). Nevertheless, [4] also proves that the result is false for $\frac{\mu}{T}$ small enough, at least in dimension 2.

Theorem 3.4. *There exists $\rho_\star \in (0,1)$ such that, for every controllable pair $(A,b) \in \mathcal{M}_2(\mathbb{R}) \times \mathbb{R}^2$ and every positive T, μ satisfying $0 < \frac{\mu}{T} < \rho_\star$, one has $RC_{\mathcal{G}}(T,\mu) < +\infty$.*

Saying that $\mathrm{RC}_{\mathcal{G}}(T,\mu) < +\infty$ means that there exists $C > 0$ such that, for every $K^{\mathrm{T}} \in \mathbb{R}^2$, one has $\mathrm{rc}_{\mathcal{G}}(T,\mu,K) \leq C$, and hence that there exists $\alpha \in \mathcal{G}(T,\mu)$ such that $\lambda^+(\alpha,K) \geq -C$. The proof of Theorem 3.4 given in [4] explicitly constructs such an α for every $K^{\mathrm{T}} \in \mathbb{R}^2$. In particular, the construction shows that, as $\|K\|$ increases, the signal α constructed oscillates faster between 0 and 1. As it is remarked in [4], one can interpret the construction by saying that the time that α passes on 1 is short enough so that the stabilizing effect of the dynamics of the system $\dot{x} = (A - bK)x$ is countered by the overshoot phenomenon occurring over small intervals of time, and it is this overshoot phenomenon that prevents the system from being stabilized at an arbitrary rate. This is only possible because α oscillates quickly between 1 and

0: in the case where $\alpha \in \mathcal{D}(T,\mu,M)$, if, for instance, α takes a certain positive value ρ at time t, the interval of time around t where α is greater than $\rho/2$ cannot be made arbitrarily small, and then we expect that the overshoot phenomenon of $\dot{x} = (A - \rho bK)x$ will eventually be countered by the stabilizing effect for K sufficiently big in norm. In other words, the argument used in the proof of the Theorem 3.4 does not apply to a signal in $\mathcal{D}(T,\mu,M)$, and hence the result $\mathrm{RC}_{\mathcal{D}}(T,\mu,M) = +\infty$ may be true, and we expect so.

This is what motivates the research of a proof of the fact that $\mathrm{RC}_{\mathcal{D}}(T,\mu,M) = +\infty$. The technique used in the proof of the Theorem 3.2 did not provide any help in this case: the direct study of a limit system comes from accelerating the dynamics of the system, which would mean that the signal α would be taken in $\mathcal{D}(T/\nu,\mu/\nu,\nu M)$ for a large constant ν and thus, in the limit $\nu \to +\infty$, the fact that α is νM-Lipschitz would provide no additional information on a weak-$\star$ limit function $\alpha_\star$. Furthermore, even if, by a change of variables, the Lipschitz continuity of α could be taken on account, the procedure of acceleration of the dynamics gives rise to a small gain K that stabilizes the system slowly. For these reasons, the research of a proof by using a limit system was considered of no direct applicability in this case.

By concentrating on the two-dimensional case of the double integrator, we could find a proof using a different technique. First of all we chose a particular form of K and a change of variables that concentrates the convergence information of the system, in such a way that we shall need only to bound the divergence rate of the solution in the new variable y in order to have convergence in the original variable x. In the new variable y, the system can be proved to rotate around the origin, and thus we can decompose the time in intervals in which the solution completes a half tour around the origin. According to the behavior of α in each interval, we are able to estimate the divergence rate of y and the estimations lead to a divergence rate that is smaller than the convergence rate given in the change of variables from x to y, which yields convergence of x at an arbitrary rate. This idea is detailed in the next section.

4 Main result

The main result we want to prove concerns the double integrator. More specifically, we fix positive constants T, μ and M with $T \geq \mu$ and we study the PEL system

$$\dot{x} = Ax + \alpha(t)bu,$$

where $x \in \mathbb{R}^2$, A is the Jordan block $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\alpha \in \mathcal{D}(T,\mu,M)$. This system can thus be written in the form

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = \alpha u, \end{cases} \qquad \alpha \in \mathcal{D}(T,\mu,M). \tag{4.1}$$

Our goal is to prove the following theorem.

Theorem 4.1. *Let T, μ and M be positive constants with $T \geq \mu$. Then, for the PEL System* (4.1)*, one has $RC_{\mathcal{D}}(T,\mu,M) = +\infty$, i.e., for every constant λ, there exists $K^T \in \mathbb{R}^2$ such that, for every $\alpha \in \mathcal{D}(T,\mu,M)$, one has $\lambda^+(\alpha,K) \leq -\lambda$.*

From now on, we suppose that T, μ, M and λ are fixed. We shall prove Theorem 4.1 by explicitly constructing the gain K that satisfies $\lambda^+(\alpha,K) \leq -\lambda$ for every $\alpha \in \mathcal{D}(T,\mu,M)$. To do so, we write $K = \begin{pmatrix} k_1 & k_2 \end{pmatrix}$ and thus the feedback $u = -Kx$ leads to the system

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -\alpha(t)k_1 & -\alpha(t)k_2 \end{pmatrix} x.$$

The variable x_1 satisfies the scalar equation

$$\ddot{x}_1 + k_2\alpha(t)\dot{x}_1 + k_1\alpha(t)x_1 = 0$$

and we have $x_2 = \dot{x}_1$.

We remark that the signal α constant and equal to 1 is in $\mathcal{D}(T,\mu,M)$, and thus a necessary condition for Theorem 4.1 to be true is that the matrix

$$A - bK = \begin{pmatrix} 0 & 1 \\ -k_1 & -k_2 \end{pmatrix}$$

is Hurwitz, which is the case if and only if $k_1 > 0$, $k_2 > 0$. In what follows, we shall restrict ourselves to search K in the form

$$K = \begin{pmatrix} k^2 & k \end{pmatrix}, \qquad k > 0. \tag{4.2}$$

The differential equation satisfied by x_1 is thus

$$\ddot{x}_1 + k\alpha(t)\dot{x}_1 + k^2\alpha(t)x_1 = 0. \tag{4.3}$$

4.1 Strategy of the proof

Let us discuss the strategy that we will use to prove Theorem 4.1. We start, in Section 4.2, by doing a change of variables on (4.3) that will help the study of the system. In addition to putting the system in a form that is easier to study and adapted to the methods that we apply latter on, this change of variables concentrate the convergence information of the system, since the original variable x and the new variable y are related by (4.5), which contains the exponential term $e^{-\frac{k}{2}\int_0^t \alpha(s)ds+\sqrt{\frac{kM}{2}}t}$, that converges to 0 as $t \to +\infty$ since α is a PE signal, and then it suffices to show that the rate of exponential growth of y is smaller than the convergence rate given by the change of variables.

We then turn to the study of the system satisfied by y in Section 4.3. We start by writing this system in polar coordinates in Section 4.3.1 and this will enable us to show in Section 4.3.2 that the solution rotates around the origin an infinity of times, which will enable us to decompose, in Section 4.3.3, the time $\mathbb{R}_+$ in the "good" intervals of $\mathcal{I}_+$, where the function γ defined in (4.7) is lower bounded by a positive constant (see Lemma 4.4), and the "bad" intervals of $\mathcal{I}_0$, where γ is small. The estimation of the growth rate of y in the intervals of $\mathcal{I}_+$ is done on Section 4.3.4: we use the fact that the polar angle θ is a strict monotone function of time to write the radial variable of the polar coordinates r in function of θ, and the a direct integration of the differential equation satisfied by $\ln r$ allows us to obtain the desired estimate. A similar technique is not possible when γ is not lower bounded by a positive constant, and then, in Section 4.3.5, we study the behavior of y in the intervals of $\mathcal{I}_0$ by using the theory of optimal control: we look for the signal γ that produces the greatest possible growth rate for y, and then, by applying Pontryagin Maximum Principle, we are able to characterize the solution y that corresponds to the maximal growth rate and finally estimate this quantity. It then suffices to put together the estimates on intervals $\mathcal{I}_0$ and $\mathcal{I}_+$ and conclude the study of y, which is done in Section 4.3.6.

Once we know the behavior of y and its growth rate, it suffices to go back to the change of variables to obtain the corresponding result in x, and this is done in Section 4.4. The estimation obtained for x shows that its convergence rate depends on k, and it then suffices to take k large enough in order to obtain the required result of convergence at an arbitrary rate, thus concluding the proof of Theorem 4.1.

4.2 Change of variables

In order to simplify the notations, we write $h = \sqrt{2kM}$. We consider the system in a new variable $y = \begin{pmatrix} y_1 & y_2 \end{pmatrix}^{\mathrm{T}}$ defined by the relations

$$\begin{cases} y_1 = x_1 e^{\frac{k}{2}\int_0^t \alpha(s)ds - \frac{h}{2}t}, \\ y_2 = \dot{y}_1 = \left(x_2 + \left(\dfrac{k}{2}\alpha(t) - \dfrac{h}{2}\right) x_1\right) e^{\frac{k}{2}\int_0^t \alpha(s)ds - \frac{h}{2}t}; \end{cases} \tag{4.4}$$

we shall justify this choice afterwards. The variables x and y are thus related by

$$y = e^{\frac{k}{2}\int_0^t \alpha(s)ds - \frac{h}{2}t} \begin{pmatrix} 1 & 0 \\ \frac{k}{2}\alpha(t) - \frac{h}{2} & 1 \end{pmatrix} x, \quad x = e^{-\frac{k}{2}\int_0^t \alpha(s)ds + \frac{h}{2}t} \begin{pmatrix} 1 & 0 \\ \frac{h}{2} - \frac{k}{2}\alpha(t) & 1 \end{pmatrix} y \tag{4.5}$$

and y_1 satisfies the differential equation

$$\ddot{y}_1 + h\dot{y}_1 + k^2\gamma(t) y_1 = 0 \tag{4.6}$$

with

$$\gamma(t) = \beta(t) + \frac{M - \dot\alpha(t)}{2k}, \qquad \beta(t) = \alpha(t)\left(1 - \tfrac{1}{4}\alpha(t)\right). \tag{4.7}$$

The system satisfied by y is

$$\dot y = \begin{pmatrix} 0 & 1 \\ -k^2\gamma(t) & -h \end{pmatrix} y. \tag{4.8}$$

Since $\alpha(t) \in [0,1]$ for every $t \in \mathbb{R}_+$, we have $\beta(t) \in [0, 3/4]$. Furthermore, since α is M-Lipschitz, β is also Lipschitz with the same Lipschitz constant, since

$$\begin{aligned} |\beta(t) - \beta(s)| &= \left|\alpha(t) - \alpha(s) - \tfrac{1}{4}\left(\alpha(t)^2 - \alpha(s)^2\right)\right| = \\ &= |\alpha(t) - \alpha(s)|\left|1 - \tfrac{\alpha(t)+\alpha(s)}{4}\right| \leq \\ &\leq |\alpha(t) - \alpha(s)| \leq M\,|t-s| \end{aligned}$$

for every $t, s \in \mathbb{R}_+$. Since α satisfies the PE condition (2.2), β satisfies

$$\int_t^{t+T} \beta(s) ds \geq \tfrac{3}{4}\mu. \tag{4.9}$$

Since $|\dot\alpha(t)| \leq M$ almost everywhere on $\mathbb{R}_+$, γ can be estimated by

$$0 \leq \gamma(t) \leq \tfrac{3}{4} + \tfrac{M}{k}$$

almost everywhere on $\mathbb{R}_+$. It also satisfies the PE condition

$$\int_t^{t+T} \gamma(s) ds \geq \tfrac{3}{4}\mu. \tag{4.10}$$

From now on, we suppose that

$$k \geq K_1(M) \tag{4.11}$$

where $K_1(M) = 4M$, so that, for almost every $t \in \mathbb{R}_+$, we have

$$0 \leq \gamma(t) \leq 1.$$

The differential equation (4.6) justifies the choice of the change of variables (4.4). In fact, the term $e^{\frac{k}{2}\int_0^t \alpha(s)ds}$ in the change of variables corresponds to a classical change of variables in second-order scalar equations (see, for instance, [5]) that eliminates the term on $\dot x_1$ from (4.3), having as effect the new term $-\frac{1}{4}k^2\alpha(t)^2 - \frac{k}{2}\dot\alpha(t)$ multiplying y_1. However, if we took only this term on the change of variables, the resulting function γ would be $\gamma(t) = \beta(t) - \frac{\dot\alpha(t)}{2k}$, which may be negative at certain times t. To apply the techniques of optimal control that we shall develop in Section 4.3.5, it is important to have a function γ that is positive almost everywhere, and that is why we introduce the term $e^{\frac{h}{2}t}$ in the change of variables. It adds the constant $\frac{kM}{2}$ multiplying y_1, which results in the fact that $\gamma(t) \geq 0$ almost everywhere on $\mathbb{R}_+$, and has also as result the new term $h\dot y_1$. We then have $\gamma(t) \geq 0$ almost everywhere as required, and the coefficient of $\dot y_1$ is no longer a time-dependent function.

Another important feature of this change of variables is that the link between the variables x and y, given by (4.5), is such that $x(t)$ behaves as $e^{-\frac{k}{2}\int_0^t \alpha(s)ds + \frac{h}{2}t} y(t)$ and, since $h = \sqrt{2kM}$ and α is persistently exciting, this exponential factor is bounded by $e^{-c_1 kt}$ for large k, for a certain positive constant c_1. We then concentrate a convergence information in the change of variables, and we no longer have to prove convergence to the origin for the system in the variables y: it is sufficient to show that the exponential growth of y is bounded by $e^{c_2 k^a t}$ for large k, for certain constants $c_2 > 0$ and $a < 1$.

This change of variables also justifies the choice of K on the form (4.2). Equation (4.6) is a linear second-order scalar differential equation, and, in the case where its coefficients are constant, $h\dot{y}_1$ can be interpreted as a damping term and $k^2\gamma y_1$ as an oscillatory term. Such a system will oscillate around the origin if $4k^2\gamma \geq h^2 = 2kM$, which is the case for k large enough. In the case where γ depends on time, the PE condition (4.10) makes one believe that a similar condition may be found in order to retrieve a certain oscillatory behavior for k large enough. This is only possible because, for k large enough, the oscillatory term in (4.6) is much larger than the damping term, which is a consequence of the choice of K in the particular form (4.2). It is thus important, in the choice (4.2), that k_1 is much larger than k_2 as k_2 increases; other kinds of choices of K in this sense would be possible. We stress that it is this oscillatory behavior that we will exploit in what follows in order to prove Theorem 4.1.

4.3 Properties of the system in the new variables

4.3.1 Polar coordinates

We now wish to study System (4.8) and the corresponding differential equation (4.6). To do so, we first write this system in polar coordinates in the plan $(y_1, \dot{y}_1)$: we define the variables $r \in \mathbb{R}_+$ and $\theta \in \mathbb{R}$ (or $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, the choice of the set where θ is taken being done according to the context) by the relations

$$\begin{aligned} r^2 &= y_1^2 + \dot{y}_1^2, \\ y_1 &= r\cos\theta, \\ \dot{y}_1 &= r\sin\theta, \end{aligned}$$

which leads to the equations

$$\dot{\theta} = -\sin^2\theta - k^2\gamma(t)\cos^2\theta - h\sin\theta\cos\theta, \tag{4.12a}$$

$$\dot{r} = r\sin\theta\cos\theta(1 - k^2\gamma(t)) - hr\sin^2\theta. \tag{4.12b}$$

Since we are dealing with a linear system, the origin is an equilibrium solution and, if we consider only the other solutions of the system, we have $r(t) > 0$ for every $t \in \mathbb{R}_+$, and thus we can write (4.12b) as

$$\frac{d}{dt}\ln r = \sin\theta\cos\theta(1 - k^2\gamma(t)) - h\sin^2\theta. \tag{4.12c}$$

4.3.2 Rotations around the origin

Let us consider Equation (4.12a). We can see that, if $\sin\theta\cos\theta \geq 0$, then $\dot{\theta} \leq 0$, and it is strictly negative except when $\sin\theta = 0$ and $\gamma(t) = 0$. If $\sin\theta\cos\theta < 0$, we still expect $\dot{\theta}$ to be "mostly" negative, meaning that, if we take k large enough, outside a certain region of the plane around the line $\cos\theta = 0$, we still have $\dot{\theta} \leq 0$, and, since h is much smaller than k^2 for k large enough, we expect that this will imply that $\lim_{t\to+\infty}\theta(t) = -\infty$, thus showing that the solution y turns clockwise (in the usual orientation of the axes y_1 and y_2) around the origin, even if, in certain points, it may go counterclockwise for a short period of time. This is the idea behind the following result.

Lemma 4.2. *There exists $K_2(T,\mu,M)$ such that, for $k > K_2(T,\mu,M)$, the solution θ of* (4.12a) *satisfies* $\lim_{t\to+\infty}\theta(t) = -\infty$.

Proof. We start by fixing $t \in \mathbb{R}_+$ and the interval $I = [t, t+T]$. Equation (4.9) shows that there exists $t_\star \in I$ such that $\beta(t_\star) \geq \frac{3\mu}{4T}$. Since β is M-Lipschitz, we have $\beta(s) \geq \frac{\mu}{2T}$ if $|s - t_\star| \leq \frac{\mu}{4MT}$, and thus, since $\gamma(s) \geq \beta(s)$, we have $\gamma(s) \geq \frac{\mu}{2T}$ for $|s - t_\star| \leq \frac{\mu}{4MT}$. If we take

$$k \geq \max\left\{1, \left(\frac{\mu}{2MT^2}\right)^4\right\}, \tag{4.13}$$

we have $\frac{\mu}{4MTk^{1/4}} \leq \frac{\mu}{4MT}$ and $\frac{\mu}{4MTk^{1/4}} \leq T/2$, which implies that at least one of the intervals $\left[t_\star - \frac{\mu}{4MTk^{1/4}}, t_\star\right]$ and $\left[t_\star, t_\star + \frac{\mu}{4MTk^{1/4}}\right]$ is included in I; let us name this interval J and write it as $J = [s_0, s_1]$, so that $s_1 - s_0 = \frac{\mu}{4MTk^{1/4}}$ and $\gamma(s) \geq \frac{\mu}{2T}$ for $s \in J$.

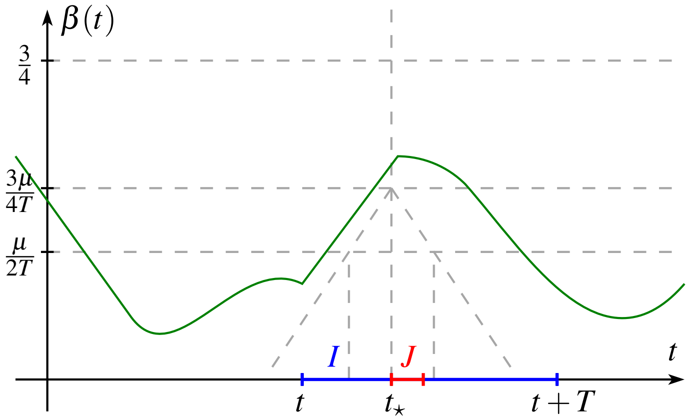

Figure 4.1: A given function β, a given interval I and the corresponding interval J. The properties we need are that $J \subset I$ and that $\gamma(s) \geq \frac{\mu}{2T}$ for $s \in J$. In this case, we can show that the solution turns clockwise around the origin in J, the number of turns increasing as $k^{3/4}$, whereas, in the rest of the interval I, the number of counterclockwise turns around the origin is an $O(k^{1/2})$, which implies that, for k large enough, $\theta(t+T) - \theta(t) \leq -2\pi$.

If $s \in J$, one can estimate $\dot\theta$ in (4.12a) by

$$-\dot\theta(s) \geq \sin^2\theta(s) + \frac{\mu k^2}{2T}\cos^2\theta(s) + h\sin\theta(s)\cos\theta(s) =$$
$$= \begin{pmatrix}\sin\theta(s) & \cos\theta(s)\end{pmatrix}\begin{pmatrix}1 & \frac{h}{2}\\ \frac{h}{2} & \frac{\mu k^2}{2T}\end{pmatrix}\begin{pmatrix}\sin\theta(s)\\ \cos\theta(s)\end{pmatrix}.$$

In particular, if

$$k > \frac{MT}{\mu}, \tag{4.14}$$

then the matrix $\begin{pmatrix}1 & \frac{h}{2}\\ \frac{h}{2} & \frac{\mu k^2}{2T}\end{pmatrix}$ is positive definite and thus $\dot\theta(s) < 0$ for every $s \in J$. Therefore θ is strictly decreasing on J and is a bijection between J and its image $\theta(J)$. One can write Equation (4.12a) on J as

$$\frac{\dot\theta}{\sin^2\theta + k^2\gamma\cos^2\theta + h\sin\theta\cos\theta} = -1 \tag{4.15}$$

and, by integrating from s_0 to s_1 and using the relation

$$\int_{-\pi/2}^{\pi/2} \frac{d\theta}{\sin^2\theta + a\cos^2\theta + b\sin\theta\cos\theta} = \frac{2\pi}{\sqrt{4a-b^2}}, \qquad a>0,\, b^2<4a,$$

which can be calculated directly by the change of variables $\hat{t} = \tan\theta$, we obtain

$$\begin{aligned} \frac{\mu}{4MTk^{1/4}} = s_1 - s_0 = -\int_{s_0}^{s_1} \frac{\dot\theta(s)}{\sin^2\theta(s) + k^2\gamma(s)\cos^2\theta(s) + h\sin\theta(s)\cos\theta(s)} ds \le \\ \le \int_{\theta(s_1)}^{\theta(s_0)} \frac{d\theta}{\sin^2\theta + \frac{k^2\mu}{2T}\cos^2\theta + h\sin\theta\cos\theta} \le \\ \le \int_{\theta(s_1)}^{\theta(s_1)+\pi(N+1)} \frac{d\theta}{\sin^2\theta + \frac{k^2\mu}{2T}\cos^2\theta + h\sin\theta\cos\theta} = \\ = \frac{2\pi(N+1)}{\sqrt{\frac{2k^2\mu}{T} - h^2}} = \frac{2\pi(N+1)}{\sqrt{\frac{2\mu}{T}k^2 - 2Mk}}, \end{aligned} \tag{4.16}$$

where N is the number of rotations of angle π done during the interval J, i.e.,

$$N = \left\lfloor \frac{\theta(s_0) - \theta(s_1)}{\pi} \right\rfloor.$$

Therefore

$$\theta(s_0) - \theta(s_1) \ge \pi N \ge k^{3/4} \frac{\mu}{8MT} \sqrt{\frac{2\mu}{T} - \frac{2M}{k}} - \pi. \tag{4.17}$$

On the other hand, one can estimate $\dot\theta$ in (4.12a) for every $s \in I$ by

$$\dot\theta(s) \le h,$$

so that

$$\theta(s_0) - \theta(t) \le h(s_0 - t), \qquad \theta(t+T) - \theta(s_1) \le h(t+T-s_1). \tag{4.18}$$

Thus, by (4.17) and (4.18), we obtain

$$\theta(t+T) - \theta(t) \le \sqrt{2kM}T - k^{3/4} \frac{\mu}{8MT} \sqrt{\frac{2\mu}{T} - \frac{2M}{k}} + \pi.$$

The expression on the right-hand side tends to $-\infty$ as $k \to +\infty$ and the parameters T, μ and M are fixed, and so there exists $K_\star(T,\mu,M)$ such that, if

$$k \ge K_\star(T,\mu,M), \tag{4.19}$$

then

$$\sqrt{2kM}T - k^{3/4} \frac{\mu}{8MT} \sqrt{\frac{2\mu}{T} - \frac{2M}{k}} + \pi \le -2\pi$$

and thus

$$\theta(t+T) - \theta(t) \le -2\pi.$$

We group conditions (4.13), (4.14) and (4.19) in a single one by setting

$$K_2(T,\mu,M) = \max\left\{1, \left(\frac{\mu}{2MT^2}\right)^4, \frac{MT}{\mu}, K_\star(T,\mu,M)\right\}$$

and asking that

$$k > K_2(T,\mu,M).$$

Under this condition, the solution completes at least one complete clockwise rotation by the end of the interval $[t,t+T]$. This result is true for every $t \in \mathbb{R}_+$ and thus an immediate induction shows that

$$\theta(t+nT) - \theta(t) \leq -2n\pi$$

for every $n \in \mathbb{N}$, so that, for every $t \in \mathbb{R}_+$,

$$\theta(t) = \theta\left(\{t/T\}T + \lfloor t/T \rfloor T\right) \leq \theta\left(\{t/T\}T\right) - 2\lfloor t/T \rfloor \pi \tag{4.20}$$

where $\{x\} = x - \lfloor x \rfloor \in [0,1)$. Since θ is bounded on the interval $[0,T]$, Inequality (4.20) shows that $\lim_{t\to+\infty}\theta(t) = -\infty$, thus completing the proof. ■

4.3.3 Decomposition of the time in intervals $\mathcal{I}_+$ and $\mathcal{I}_0$

Using Lemma 4.2, we can decompose $\mathbb{R}_+$ in a sequence of intervals (depending on α) on which the solution rotates by an angle π around the origin. More precisely, we define the sequence $(t_n)_{n\in\mathbb{N}}$ by induction as

$$\begin{aligned} t_0 &= \inf\{t \geq 0 \mid \tfrac{\theta(t)}{\pi} \in \mathbb{Z}\}, \\ t_n &= \inf\{t \geq t_{n-1} \mid \theta(t) = \theta(t_{n-1}) - \pi\}, \quad n \geq 1, \end{aligned} \tag{4.21}$$

and the continuity of θ and Lemma 4.2 show that this sequence is well defined. We also define the sequence of intervals $(I_n)_{n\in\mathbb{N}}$ by $I_n = [t_{n-1},t_n]$ for $n \geq 1$ and $I_0 = [0,t_0]$. The construction means thus that we wait until the solution passes through the axis y_1 for the first time and, from this moment, we divide the time in intervals on which the solution rotates by an angle π around the origin, going back to the axis y_1.

We can show a first result about the behavior of θ on these intervals.

Lemma 4.3. *Let $n \geq 1$. Then, for every $t \in I_n = [t_{n-1},t_n]$, one has*

$$\theta(t_n) \leq \theta(t) \leq \theta(t_{n-1}). \tag{4.22}$$

Proof. The first inequality in (4.22) is a consequence of the definition of t_n: if there was $t \in I_n$ with $\theta(t) < \theta(t_n)$, then, by the continuity of θ, there would be $s \in]t_{n-1},t[$ such that $\theta(s) = \theta(t_n) = \theta(t_{n-1}) - \pi$, and thus, by definition of t_n, we would have $t_n \leq s < t < t_n$, which is a contradiction, and thus we have $\theta(t) \geq \theta(t_n)$ for every $t \in I_n$.

The second inequality in (4.22) can also be proved by contradiction. We suppose that there exists $t \in I_n$ such that $\theta(t) > \theta(t_{n-1})$. Then, by continuity of θ, there exists $s_0, s_1 \in [t_{n-1},t]$ such that $\theta(s_0) = \theta(t_{n-1})$, $\theta(s_1) > \theta(t_{n-1})$ and $\theta(s) \in [\theta(t_{n-1}), \theta(t_{n-1}) + \pi/2]$ for every $s \in [s_0,s_1]$. But we have that $\theta(t_{n-1}) = 0 \mod \pi$, so that $\sin\vartheta\cos\vartheta \geq 0$ for $\vartheta \in [\theta(t_{n-1}), \theta(t_{n-1}) + \pi/2]$, and thus, by (4.12a), $\dot\theta(s) \leq 0$ for almost every $s \in [s_0,s_1]$, which contradicts the fact that $\theta(s_0) = \theta(t_{n-1})$ and $\theta(s_1) > \theta(t_{n-1})$ since θ is absolutely continuous. We then have $\theta(t) \leq \theta(t_{n-1})$ for every $t \in I_n$. ■

We now split the intervals of the sequence $(I_n)_{n\in\mathbb{N}^*}$ into two classes, $\mathcal{I}_+$ and $\mathcal{I}_0$, according to the behavior of β on these intervals. We define

$$\begin{aligned} \mathcal{I}_+ &= \left\{I_n \mid n \in \mathbb{N}^*, \exists t \in I_n, \beta(t) \geq 2/\sqrt{k}\right\}, \\ \mathcal{I}_0 &= \left\{I_n \mid n \in \mathbb{N}^*, \forall t \in I_n, \beta(t) < 2/\sqrt{k}\right\}. \end{aligned}$$

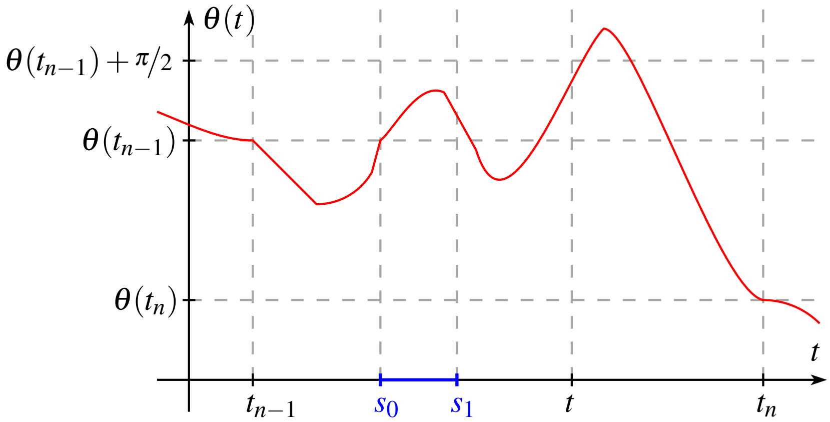

Figure 4.2: Contradiction used to prove the second inequality in (4.22). The existence of t such that $\theta(t) > \theta(t_{n-1})$ makes it possible to construct an interval $[s_0, s_1]$ where $\theta(s_1) > \theta(s_0)$ but $\dot{\theta} \leq 0$, thus leading to a contradiction.

4.3.4 Estimations on intervals $\mathcal{I}_+$

We start by studying the intervals in the class $\mathcal{I}_+$. We first claim that, for k large enough, we have $\gamma(t) \geq 1/\sqrt{k}$ for almost every $t \in I$ and every $I \in \mathcal{I}_+$.

Lemma 4.4. *There exists $K_3(M)$ such that, for $k > K_3(M)$ and for every $I \in \mathcal{I}_+$, one has $\beta(t) \geq 1/\sqrt{k}$ for every $t \in I$ and $\gamma(t) \geq 1/\sqrt{k}$ for almost every $t \in I$.*

Proof. We fix an interval $I = [t_{n-1}, t_n] \in \mathcal{I}_+$ and we note $t_\star \in I$ an element of I such that $\beta(t_\star) \geq 2/\sqrt{k}$. Since β is M-Lipschitz, for every t such that $|t - t_\star| \leq \frac{1}{M\sqrt{k}}$, we have $1/\sqrt{k} \leq \beta(t) \leq 3/\sqrt{k}$. In particular, since $\gamma(t) \geq \beta(t)$ on $\mathbb{R}_+$, we have $\gamma(t) \geq 1/\sqrt{k}$ for $|t - t_\star| \leq \frac{1}{M\sqrt{k}}$.

The idea is to show that, for k large enough, one must have $I \subset \left[t_\star - \frac{1}{M\sqrt{k}}, t_\star + \frac{1}{M\sqrt{k}}\right]$, which we do by saying that, for k large enough, the number of rotations of angle π around the origin done on each of the intervals $\left[t_\star - \frac{1}{M\sqrt{k}}, t_\star\right]$ and $\left[t_\star, t_\star + \frac{1}{M\sqrt{k}}\right]$ is larger than 1, which is the number of rotations of angle π around the origin done on I.

We take $s_0, s_1 \in \left[t_\star - \frac{1}{M\sqrt{k}}, t_\star + \frac{1}{M\sqrt{k}}\right]$, $s_0 < s_1$. For every $s \in [s_0, s_1]$, we have

$$-\dot{\theta}(s) \geq \sin^2\theta(s) + k^{3/2}\cos^2\theta(s) + h\sin\theta(s)\cos\theta(s) =$$
$$= \begin{pmatrix}\sin\theta(s) & \cos\theta(s)\end{pmatrix} \begin{pmatrix} 1 & \frac{h}{2} \\ \frac{h}{2} & k^{3/2}\end{pmatrix} \begin{pmatrix}\sin\theta(s) \\ \cos\theta(s)\end{pmatrix},$$

and the matrix $\begin{pmatrix} 1 & \frac{h}{2} \\ \frac{h}{2} & k^{3/2}\end{pmatrix}$ is positive definite if

$$k > \frac{M^2}{4}. \tag{4.23}$$

We take k satisfying (4.23). We can thus write Equation (4.12a) on $[s_0, s_1]$ as (4.15), and by integrating as in (4.16), we obtain

$$s_1 - s_0 \leq \int_{\theta(s_1)}^{\theta(s_1)+\pi(N(s_0,s_1)+1)} \frac{d\theta}{\sin^2\theta + k^{3/2}\cos^2\theta + h\sin\theta\cos\theta} =$$
$$= \frac{2\pi(N(s_0,s_1)+1)}{\sqrt{4k^{3/2} - 2Mk}} = \frac{\pi(N(s_0,s_1)+1)}{k^{3/4}\sqrt{1 - \frac{M}{2k^{1/2}}}},$$

where

$$N(s_0,s_1)=\left\lfloor\frac{\theta(s_0)-\theta(s_1)}{\pi}\right\rfloor$$

is the number of rotations of angle π around the origin done by the solution between s_0 and s_1. Hence

$$N(s_0,s_1)\geq k^{3/4}\frac{(s_1-s_0)}{\pi}\sqrt{1-\frac{M}{2k^{1/2}}}-1,$$

and, in particular,

$$N\left(t_\star-\tfrac{1}{M\sqrt{k}},t_\star\right)\geq\frac{k^{1/4}}{M\pi}\sqrt{1-\frac{M}{2k^{1/2}}}-1,$$
$$N\left(t_\star,t_\star+\tfrac{1}{M\sqrt{k}}\right)\geq\frac{k^{1/4}}{M\pi}\sqrt{1-\frac{M}{2k^{1/2}}}-1.$$

For M fixed, we have $\frac{k^{1/4}}{M\pi}\sqrt{1-\frac{M}{2k^{1/2}}}-1\xrightarrow[k\to+\infty]{}+\infty$, and thus there exists $K_\star(M)$ such that, for

$$k>K_\star(M),\tag{4.24}$$

one has

$$\frac{k^{1/4}}{M\pi}\sqrt{1-\frac{M}{2k^{1/2}}}-1>1.$$

Therefore

$$N\left(t_\star-\tfrac{1}{M\sqrt{k}},t_\star\right)>1,\qquad N\left(t_\star,t_\star+\tfrac{1}{M\sqrt{k}}\right)>1,$$

and then

$$\theta(t_\star)-\theta\left(t_\star+\tfrac{1}{M\sqrt{k}}\right)>\pi,\qquad\theta\left(t_\star-\tfrac{1}{M\sqrt{k}}\right)-\theta(t_\star)>\pi.\tag{4.25}$$

By definition of I, we have $\theta(t_{n-1})-\theta(t_n)=\pi$, and, by Lemma 4.3, $\theta(t_n)\leq\theta(t)\leq\theta(t_{n-1})$ for every $t\in I$; the fact that $t_\star\in I$ and (4.25) show that $t_\star-\frac{1}{M\sqrt{k}}\notin I$, $t_\star+\frac{1}{M\sqrt{k}}\notin I$, from where we conclude that

$$t_\star-\tfrac{1}{M\sqrt{k}}<t_{n-1},\qquad t_\star+\tfrac{1}{M\sqrt{k}}>t_n,$$

and then $I\subset\left[t_\star-\frac{1}{M\sqrt{k}},t_\star+\frac{1}{M\sqrt{k}}\right]$. We now group (4.23) and (4.24) by setting

$$K_3(M)=\max\left\{\frac{M^2}{4},K_\star(M)\right\}$$

and asking that

$$k>K_3(M).$$

Under this hypothesis, we have $I\subset\left[t_\star-\frac{1}{M\sqrt{k}},t_\star+\frac{1}{M\sqrt{k}}\right]$ and, since $\beta(t)\geq 1/\sqrt{k}$ for every t such that $|t-t_\star|\leq\frac{1}{M\sqrt{k}}$ and $\gamma(t)\geq 1/\sqrt{k}$ for almost every t such that $|t-t_\star|\leq\frac{1}{M\sqrt{k}}$, we obtain the desired result. ■

By using this result, we can estimate the divergence rate of the solution over an interval on the class $\mathcal{I}_+$.

Lemma 4.5. *There exists $K_4(M)$ such that, for $k > K_4(M)$ and for every $I = [t_{n-1}, t_n] \in \mathcal{I}_+$, the solution of* (4.12c) *satisfies*

$$r(t_n) \leq r(t_{n-1}) e^{4Mk^{1/2}(t_n - t_{n-1})}. \tag{4.26}$$

Proof. We start by taking

$$k > K_3(M), \tag{4.27}$$

so that we can apply Lemma 4.4 and obtain that $\beta(t) \geq 1/\sqrt{k}$ for every $t \in I$ and $\gamma(t) \geq 1/\sqrt{k}$ for almost every $t \in I$. We thus have, for $t \in I$,

$$\begin{aligned} -\dot{\theta}(t) &\geq \sin^2\theta(t) + k^{3/2}\cos^2\theta(t) + h\sin\theta(t)\cos\theta(t) = \\ &= \begin{pmatrix} \sin\theta(t) & \cos\theta(t) \end{pmatrix} \begin{pmatrix} 1 & \frac{h}{2} \\ \frac{h}{2} & k^{3/2} \end{pmatrix} \begin{pmatrix} \sin\theta(t) \\ \cos\theta(t) \end{pmatrix} > 0 \end{aligned}$$

for, since $k > K_3(M)$, we have in particular (4.23) and thus the above matrix is positive definite. Hence θ is a continuous strictly decreasing function on I, being thus a bijection between $I = [t_{n-1}, t_n]$ and its image $[\theta(t_n), \theta(t_{n-1})]$. We note by τ the inverse of θ, defined on $[\theta(t_n), \theta(t_{n-1})]$; τ thus satisfies

$$\frac{d\tau}{d\vartheta}(\vartheta) = \frac{1}{\dot{\theta}(\tau(\vartheta))} = -\frac{1}{\sin^2\vartheta + k^2\gamma(\tau(\vartheta))\cos^2\vartheta + h\sin\vartheta\cos\vartheta}. \tag{4.28}$$

We note $\rho = r \circ \tau$, and hence, by using Equations (4.12c) and (4.28), we have

$$\frac{d}{d\vartheta}\ln\rho = -\frac{\sin\vartheta\cos\vartheta(1 - k^2\gamma\circ\tau(\vartheta)) - h\sin^2\vartheta}{\sin^2\vartheta + k^2\gamma\circ\tau(\vartheta)\cos^2\vartheta + h\sin\vartheta\cos\vartheta}.$$

We can integrate this expression from $\theta(t_n)$ to $\theta(t_{n-1}) = \theta(t_n) + \pi$, obtaining

$$\ln\frac{r(t_n)}{r(t_{n-1})} = \int_{\theta(t_n)}^{\theta(t_n)+\pi} F(\vartheta, \gamma\circ\tau(\vartheta))d\vartheta$$

with

$$F(\vartheta, \gamma) = \frac{\sin\vartheta\cos\vartheta(1 - k^2\gamma) - h\sin^2\vartheta}{\sin^2\vartheta + k^2\gamma\cos^2\vartheta + h\sin\vartheta\cos\vartheta}.$$

If $\gamma_0 \geq 1/\sqrt{k}$ is constant, then

$$\int_{\theta(t_n)}^{\theta(t_n)+\pi} F(\vartheta, \gamma_0)d\vartheta \leq 0; \tag{4.29}$$

to see so, it suffices to see that, since F is π-periodic on ϑ, this integral can be taken in any interval of length π, and that, by making the change of variables given by $\hat{t} = \tan\vartheta$, one has

$$\begin{aligned} \int_{-\pi/2}^{\pi/2} F(\vartheta, \gamma_0)d\vartheta &= \int_{-\infty}^{+\infty} \frac{(1 - k^2\gamma_0)\hat{t} - h\hat{t}^2}{(\hat{t}^2 + h\hat{t} + k^2\gamma_0)(\hat{t}^2 + 1)} d\hat{t} \leq \\ &\leq \int_{-\infty}^{+\infty} \frac{(1 - k^2\gamma_0)\hat{t}}{(a_0\hat{t}^2 + b_0)(\hat{t}^2 + 1)} d\hat{t} = 0 \end{aligned}$$

with $a_0 = \frac{k^2\gamma_0 - h^2/4}{k^2\gamma_0 + h^2/4}$, $b_0 = \frac{k^2\gamma_0}{2} - \frac{h^2}{8}$; these are both positive since $\gamma_0 \geq 1/\sqrt{k}$ and k satisfies (4.23), and they are chosen in such a way that $\hat{t}^2 + h\hat{t} + k^2\gamma_0 \geq a_0\hat{t}^2 + b_0$ for every $\hat{t} \in \mathbb{R}$.

By (4.29), we have

$$\ln \frac{r(t_n)}{r(t_{n-1})} \leq \int_{\theta(t_n)}^{\theta(t_n)+\pi} [F(\vartheta, \gamma \circ \tau(\vartheta)) - F(\vartheta, \gamma_0)]\, d\vartheta. \tag{4.30}$$

We compute

$$\frac{\partial F}{\partial \gamma}(\vartheta, \gamma) = -\frac{k^2 \sin\vartheta \cos\vartheta}{(\sin^2\vartheta + k^2\gamma \cos^2\vartheta + h \sin\vartheta\cos\vartheta)^2},$$

and thus, for $t \in I$,

$$\left| \frac{\partial F}{\partial \gamma}(\vartheta, \gamma(t)) \right| \leq \frac{k^2 |\sin\vartheta|\,|\cos\vartheta|}{(\sin^2\vartheta + k^{3/2}\cos^2\vartheta + h\sin\vartheta\cos\vartheta)^2}.$$

We now take $\gamma_0 = \beta(t_{n-1})$ in (4.30), obtaining

$$\ln \frac{r(t_n)}{r(t_{n-1})} \leq \int_{\theta(t_n)}^{\theta(t_n)+\pi} \frac{k^2 |\sin\vartheta|\,|\cos\vartheta|}{(\sin^2\vartheta + k^{3/2}\cos^2\vartheta + h\sin\vartheta\cos\vartheta)^2} |\gamma \circ \tau(\vartheta) - \beta(t_{n-1})|\, d\vartheta. \tag{4.31}$$

For almost every $t \in I$, one can estimate

$$|\gamma(t) - \beta(t_{n-1})| \leq |\beta(t) - \beta(t_{n-1})| + \left| \frac{\dot\alpha(t)}{2k} \right| \leq M(t_n - t_{n-1}) + \frac{M}{2k}. \tag{4.32}$$

We take k satisfying (4.11), which means that $0 \leq \gamma(t) \leq 1$ for almost every $t \in \mathbb{R}_+$, and thus, by integrating (4.15) from t_{n-1} to t_n, we obtain

$$\begin{aligned} t_n - t_{n-1} &= -\int_{t_{n-1}}^{t_n} \frac{\dot\theta(s)}{\sin^2\theta(s) + k^2\gamma(s)\cos^2\theta(s) + h\sin\theta(s)\cos\theta(s)} ds \geq \\ &\geq \int_{\theta(t_n)}^{\theta(t_n)+\pi} \frac{d\theta}{\sin^2\theta + k^2\cos^2\theta + h\sin\theta\cos\theta} = \frac{\pi}{k\sqrt{1-\frac{M}{2k}}}, \end{aligned}$$

from where we get

$$\frac{1}{k} \leq \frac{t_n - t_{n-1}}{\pi}$$

and thus (4.32) becomes

$$|\gamma(t) - \beta(t_{n-1})| \leq M\left(1 + \tfrac{1}{2\pi}\right)(t_n - t_{n-1}) < 2M(t_n - t_{n-1}).$$

We use this estimate in (4.31), which leads to

$$\ln \frac{r(t_n)}{r(t_{n-1})} \leq 2k^2 M(t_n - t_{n-1}) \int_{\theta(t_n)}^{\theta(t_n)+\pi} \frac{|\sin\vartheta|\,|\cos\vartheta|}{(\sin^2\vartheta + k^{3/2}\cos^2\vartheta + h\sin\vartheta\cos\vartheta)^2} d\vartheta. \tag{4.33}$$

In order to calculate the integral in (4.33), we use the π-periodicity of the integrand and that, for $a > 0$ and $b^2 < 4a$, we have

$$\int_{-\pi/2}^{\pi/2} \frac{|\sin\vartheta|\,|\cos\vartheta|}{(\sin^2\vartheta + a\cos^2\vartheta + b\sin\vartheta\cos\vartheta)^2} d\vartheta = \frac{1}{A} + \frac{B}{A^{3/2}} \arctan\left(B/\sqrt{A}\right) \leq \frac{1}{A}\left(1 + \frac{\pi}{2}C\right)$$

with $A = a - {}^{b^2}\!/_4 > 0$, $B = {}^{b}\!/_2$ and $C = {}^{B}\!/\!\sqrt{A} = \frac{b}{\sqrt{4a-b^2}}$. Applying this to (4.33) gives

$$\ln \frac{r(t_n)}{r(t_{n-1})} \leq \frac{2k^{1/2}M(t_n - t_{n-1})}{1 - \frac{M}{2k^{1/2}}} \left(1 + \frac{\pi}{2}\sqrt{\frac{kM}{2k^{3/2} - kM}}\right)$$

and, since $\frac{1}{1-\frac{M}{2k^{1/2}}}\left(1+\frac{\pi}{2}\sqrt{\frac{kM}{2k^{3/2}-kM}}\right) \xrightarrow[k\to+\infty]{} 1$, there exists $K_\star(M)$ such that, if

$$k \geq K_\star(M), \tag{4.34}$$

then $\frac{1}{1-\frac{M}{2k^{1/2}}}\left(1+\frac{\pi}{2}\sqrt{\frac{kM}{2k^{3/2}-kM}}\right) \leq 2$, and thus

$$\ln \frac{r(t_n)}{r(t_{n-1})} \leq 4k^{1/2}M(t_n - t_{n-1}).$$

We group hypothesis (4.11), (4.27) and (4.34) done on k by setting

$$K_4(M) = \max\{K_1(M), K_3(M), K_\star(M)\}$$

and requiring that

$$k > K_4(M).$$

Under this hypothesis, we obtain

$$r(t_n) \leq r(t_{n-1})e^{4Mk^{1/2}(t_n - t_{n-1})}.$$

■

4.3.5 Estimations on intervals $\mathcal{I}_0$

Lemma 4.5 allows us to estimate the growth of the norm after a rotation of angle π in an interval of the class $\mathcal{I}_+$. We now wish to obtain a similar result for the intervals of the class $\mathcal{I}_0$; to do so, we start by obtaining a first result characterizing the duration of these intervals and the behavior of γ.

Lemma 4.6. *There exists $K_5(T,\mu,M)$ such that, if $k > K_5(T,\mu,M)$, then for every $I = [t_{n-1}, t_n] \in \mathcal{I}_0$ one has $\gamma(t) \leq {}^{3}\!/\!\sqrt{k}$ for almost every $t \in I$ and*

$$\frac{\pi}{1 + h + 3k^{3/2}} \leq t_n - t_{n-1} < T.$$

Proof. We fix $I = [t_{n-1}, t_n] \in \mathcal{I}_0$. If

$$k \geq M^2, \tag{4.35}$$

then $0 \leq \gamma(t) - \beta(t) \leq \frac{M}{k} \leq \frac{1}{\sqrt{k}}$, and thus $\gamma(t) \leq {}^{3}\!/\!\sqrt{k}$ almost everywhere on I. Also, if

$$k > \left(\frac{8T}{3\mu}\right)^2, \tag{4.36}$$

we have $\beta(t) < 2/\sqrt{k} < \frac{3\mu}{4T}$, and thus, by the persistence of excitation (4.9) of β, we obtain that $t_n - t_{n-1} < T$. Furthermore, by (4.12a), we obtain $-\dot{\theta} \leq 1 + 3k^{3/2} + h$ almost everywhere on I, and then by integrating on I we deduce that $t_n - t_{n-1} \geq \frac{\pi}{1+h+3k^{3/2}}$. So, by defining

$$K_5(T,\mu,M) = \max\left\{M^2, \left(\frac{8T}{3\mu}\right)^2\right\},$$

inequalities (4.35) and (4.36) are satisfied if

$$k > K_5(T,\mu,M),$$

thus giving the desired result. ■

We suppose from now on that $k > K_5(T,\mu,M)$. Our goal now is to obtain a result similar to Lemma 4.5 for the case of an interval $I \in \mathcal{I}_0$. Le us start by defining the class $\mathfrak{D}(T,\mu,M,k)$ where we take γ.

Definition 4.7. We define the class $\mathfrak{D}(T,\mu,M,k)$ as

$$\mathfrak{D}(T,\mu,M,k) = \left\{\alpha\left(1 - \tfrac{1}{4}\alpha\right) + \frac{M-\dot{\alpha}}{2k},\ \alpha \in \mathcal{D}(T,\mu,M)\right\}.$$

We fix $I = [t_{n-1}, t_n] \in \mathcal{I}_0$. We remark that, if $\gamma \in \mathfrak{D}(T,\mu,M,k)$, then, for every $t_0 \in \mathbb{R}_+$, the function $t \mapsto \gamma(t+t_0)$ is also in $\mathfrak{D}(T,\mu,M,k)$. Up to a translation in time, we can then suppose $I = [0,\tau]$ with $\tau = t_n - t_{n-1} \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right)$. The solution $r(\tau)$ of (4.12c) at time τ can be written as

$$r(\tau) = r(0)e^{\Lambda\tau}$$

for a certain constant Λ. We know by construction that $r(0)$ is in the axis y_1 and thus, since System (4.8) is linear, we conclude by homogeneity that Λ does not depend on the particular value of $r(0)$, depending only on τ and $r(\tau)$. Our goal is to estimate Λ uniformly with respect to the class of signals $\gamma \in \mathfrak{D}(T,\mu,M,k)$ and with respect to all intervals $I \in \mathcal{I}_0$ for a given choice of γ. We can thus estimate Λ by the maximal value of $\frac{1}{\tau}\ln\frac{\|y(\tau)\|}{\|y(0)\|}$ over all $\tau \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right)$ and all $\gamma \in \mathfrak{D}(T,\mu,M,k)$ with $\gamma(t) < 3/\sqrt{k}$, where y is a solution of (4.8) with both $y(0)$ and $y(\tau)$ in the axis y_1. That is, Λ is estimated by the solution of the problem

$$\begin{cases} \text{Find } \sup \dfrac{1}{\tau}\ln\dfrac{\|y(\tau)\|}{\|y(0)\|} \text{ with} \\ \tau \in \left[\dfrac{\pi}{1+h+3k^{3/2}}, T\right], \quad \gamma \in \mathfrak{D}(T,\mu,M,k), \quad \gamma(t) < 3/\sqrt{k} \text{ on } [0,\tau], \\ \dot{y} = \begin{pmatrix} 0 & 1 \\ -k^2\gamma(t) & -h \end{pmatrix} y, \quad y(0) = \begin{pmatrix} y_1(0) \\ 0 \end{pmatrix}, \quad y(\tau) = \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \\ y_1(0), \xi \in \mathbb{R}^*, \quad y_1(0)\xi < 0. \end{cases} \tag{4.37}$$

We can choose $y_1(0) = -1$ without loss of generality since the equation satisfied by y is linear. We can also see that, by enlarging the class where we take γ and taking $\gamma \in L^\infty([0,\tau],[0,3/\sqrt{\gamma}])$, we obtain a problem whose solution is bigger than the solution of (4.37), and thus Λ is also estimated by the solution of the problem

$$\begin{cases} \text{Find } \sup \frac{1}{\tau} \ln \|y(\tau)\| \text{ with} \\ \tau \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right], \quad I = [0,\tau], \quad \gamma \in L^\infty(I,[0,1]), \\ \dot{y} = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma(t) & -h \end{pmatrix} y, \quad y(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad y(\tau) \in \left\{ \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \xi \in \mathbb{R}_+ \right\}. \end{cases} \tag{4.38}$$

We now precise and summarize all the preceding discussion in the following result.

Lemma 4.8. *Let $\mathbf{\Lambda}(T,M,k)$ be the solution of Problem* (4.38) *and take $K_5(T,\mu,M)$ as in Lemma 4.6. If $k > K_5(T,\mu,M)$, then, for every $\gamma \in \mathfrak{D}(T,\mu,M,k)$ and for every $I = [t_{n-1},t_n] \in \mathfrak{I}_0$, we have*

$$r(t_n) \leq r(t_{n-1}) e^{\mathbf{\Lambda}(T,M,k)(t_n - t_{n-1})}. \tag{4.39}$$

Proof. Fix $\gamma \in \mathfrak{D}(T,\mu,M,k)$ and $I = [t_{n-1},t_n] \in \mathfrak{I}_0$. We take $k > K_5(T,\mu,M)$ in order to apply Lemma 4.6. We define $\tau = t_n - t_{n-1}$, and thus Lemma 4.6 shows that $\tau \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right)$ and $\gamma(t) \leq 3/\sqrt{k}$ for almost every $t \in I$.

We note $\overline{\gamma}(t) = \frac{\sqrt{k}}{3}\gamma(t+t_{n-1})$ for every $t \in I$, and thus $\overline{\gamma} \in L^\infty(\overline{I},[0,1])$ with $\overline{I} = [0,\tau]$. We note by y the solution of (4.8) with a non-zero initial condition and by z the function defined by $z(t) = -\frac{\operatorname{sign}(y_1(t_{n-1}))}{\|y(t_{n-1})\|} y(t+t_{n-1})$. We see that z is well defined since $\|y(t_{n-1})\| \neq 0$, and z satisfies

$$\dot{z} = \begin{pmatrix} 0 & 1 \\ -k^2\gamma(t+t_{n-1}) & -h \end{pmatrix} z = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\overline{\gamma}(t) & -h \end{pmatrix} z.$$

By definition of I, both $y(t_{n-1})$ and $y(t_n)$ are on the axis y_1, on opposite sides of the origin, and thus both $z(0)$ and $z(\tau)$ are on the axis z_1 on opposite sides of the origin; by the definition of z, we can thus write

$$z(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad z(\tau) \in \left\{ \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \xi \in \mathbb{R}_+^* \right\}.$$

It suffices now to see that, by the definition of $\mathbf{\Lambda}(T,M,k)$, one has

$$\frac{1}{\tau} \ln \|z(\tau)\| \leq \mathbf{\Lambda}(T,M,k),$$

and thus

$$\|z(\tau)\| \leq e^{\mathbf{\Lambda}(T,M,k)\tau}.$$

By the definition of z and τ, we obtain (4.39). ■

We can now concentrate on the problem of solving (4.38). We start by proving that the sup in this problem is attained.

Lemma 4.9. *Let $k > K_5(T,\mu,M)$ where K_5 is defined in Lemma 4.6 and we note $\mathbf{\Lambda}(T,M,k)$ the solution of Problem* (4.38). *Then there exist $\tau_\star \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right]$ and $\gamma_\star \in L^\infty(I_\star,[0,1])$, where $I_\star = [0,\tau_\star]$, such that, if $y_\star$ denotes the solution of*

$$\dot{y}_\star = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma_\star(t) & -h \end{pmatrix} y_\star, \quad y_\star(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

then

$$y_\star(\tau) \in \left\{ \begin{pmatrix} \xi \\ 0 \end{pmatrix},\ \xi \in \mathbb{R}_+ \right\}$$

and

$$\frac{1}{\tau_\star} \ln \|y_\star(\tau_\star)\| = \mathbf{\Lambda}(T,M,k).$$

Proof. We start by taking a sequence $(\tau_n,\gamma_n)_{n\in\mathbb{N}}$ with $\tau_n \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right]$, $I_n = [0,\tau_n]$ and $\gamma_n \in L^\infty(I_n,[0,1])$, such that, by denoting y_n the solution of

$$\begin{cases} \dot{y}_n = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma_n(t) & -h \end{pmatrix} y_n, \\ y_n(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \end{cases} \tag{4.40}$$

we have

$$\lim_{n\to+\infty} \frac{1}{\tau_n} \ln \|y_n(\tau_n)\| = \mathbf{\Lambda}(T,M,k);$$

such a maximizing sequence exists by the definition of sup. Up to defining γ_n as 0 outside I_n, we can suppose that $\gamma_n \in L^\infty(I,[0,1])$ where $I=[0,T]$ and thus, by weak-$\star$ compactness of this space and by the compactness of $\left[\frac{\pi}{1+h+3k^{3/2}}, T\right]$, we can find a subsequence of $(\gamma_n)_{n\in\mathbb{N}}$ converging weak-$\star$ to a certain function $\gamma_\star \in L^\infty(I,[0,1])$ and such that the corresponding subsequence of $(\tau_n)_{n\in\mathbb{N}}$ converges to $\tau_\star \in \left[\frac{\pi}{1+h+3k^{3/2}}, T\right]$; to simplify the notation, we still write $(\gamma_n)_{n\in\mathbb{N}}$ and $(\tau_n)_{n\in\mathbb{N}}$ for these subsequences.

We remark that $\gamma_\star$ is equal 0 almost everywhere outside $I_\star = [0,\tau_\star]$ since, for every function $\varphi \in L^1([\tau_\star,T])$, one has

$$\int_{\tau_\star}^T \gamma_\star(t)\varphi(t)dt = \lim_{n\to+\infty} \int_{\tau_\star}^T \gamma_n(t)\varphi(t)dt$$

and

$$\left| \int_{\tau_\star}^T \gamma_n(t)\varphi(t)dt \right| = \begin{cases} 0 & \text{if } \tau_n \le \tau_\star, \\ \left| \int_{\tau_\star}^{\tau_n} \gamma_n(t)\varphi(t)dt \right| \le \int_{\tau_\star}^{\tau_n} |\varphi(t)|\,dt \xrightarrow[n\to+\infty]{} 0 & \text{if } \tau_n > \tau_\star. \end{cases}$$

We consider $\gamma_\star \in L^\infty(I_\star,[0,1])$. We note by $y_\star$ the solution corresponding to $\gamma_\star$, i.e., the solution of

$$\begin{cases} \dot{y}_\star = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma_\star(t) & -h \end{pmatrix} y_\star, \\ y_\star(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \end{cases} \tag{4.41}$$

By defining γ_n and $\gamma_\star$ to be 0 on $[0,T]$ outside their respective definition intervals I_n and $I_\star$, we can consider the solutions y_n and $y_\star$ of (4.40) and (4.41) to be defined on $[0,T]$ and, in this case, up to extracting a subsequence, we have $\lim_{n\to+\infty} y_n = y_\star$ uniformly on $[0,T]$. In fact, let us note $e_n = y_n - y_\star$. We write

$$A_n(t) = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma_n(t) & -h \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ -3k^{3/2} & 0 \end{pmatrix}.$$

The function e_n satisfies

$$\begin{cases} \dot{e}_n(t) = A_n(t)e_n(t) + (\gamma_n(t) - \gamma_\star(t))By_\star(t), \\ e_n(0) = (0,0)^{\mathrm{T}}, \end{cases}$$

and, by integrating this equation, we get

$$e_n(t) = \int_0^t A_n(s)e_n(s)ds + h_n(t), \qquad h_n(t) = \int_0^t (\gamma_n(s) - \gamma_\star(s))By_\star(s)ds. \tag{4.42}$$

We now apply Gronwall's Lemma to $\|e_n(t)\|$, obtaining

$$\|e_n(t)\| \leq \|h_n(t)\| + \int_0^t \|h_n(s)\| \, \|A_n(s)\| \, e^{\int_s^t \|A_n(s')\| ds'} ds. \tag{4.43}$$

If t is fixed, the weak-$\star$ convergence of γ_n to $\gamma_\star$ shows that $\lim_{n\to+\infty} h_n(t) = 0$ for every $t \in [0,T]$ and, furthermore, the sequence $(h_n)_{n\in\mathbb{N}}$ is uniformly bounded on $[0,T]$, which shows that, by the Dominated Convergence Theorem,

$$\lim_{n\to+\infty} \int_0^t \|h_n(s)\| \, \|A_n(s)\| \, e^{\int_s^t \|A_n(s')\| ds'} ds = 0$$

for every $t \in [0,T]$, since $(\|A_n\|)_{n\in\mathbb{N}}$ is also uniformly bounded. Thus $\lim_{n\to+\infty} e_n(t) = 0$ for every $t \in [0,T]$. Since $(h_n)_{n\in\mathbb{N}}$ is uniformly bounded, $(e_n)_{n\in\mathbb{N}}$ is also because of (4.43), and (4.42) shows that, for $t > t'$,

$$e_n(t) - e_n(t') = \int_{t'}^t A_n(s)e_n(s)ds + \int_{t'}^t (\gamma_n(s) - \gamma_\star(s))By_\star(s)ds,$$

which, with the uniform bound of $(e_n)_{n\in\mathbb{N}}$, shows that this sequence is equicontinuous. Hence, by Arzelà-Ascoli Theorem, up to taking a subsequence, $(e_n)_{n\in\mathbb{N}}$ converges uniformly and, since it converges punctually to 0, its uniform limit is the function 0, which shows that $\lim_{n\to+\infty} y_n = y_\star$ uniformly on $[0,T]$.

The uniform convergence enables us to show the conclusion of the lemma. In fact, since $y_n(\tau_n) \in \left\{ \begin{pmatrix} \xi & 0 \end{pmatrix}^{\mathrm{T}}, \, \xi \in \mathbb{R}_+ \right\}$ and $\lim_{n\to+\infty} y_n(\tau_n) = y_\star(\tau_\star)$ by uniform convergence, we conclude that $y_\star(\tau_\star) \in \left\{ \begin{pmatrix} \xi & 0 \end{pmatrix}^{\mathrm{T}}, \, \xi \in \mathbb{R}_+ \right\}$ since this set is closed. The uniform convergence of y_n to $y_\star$ also shows that

$$\frac{1}{\tau_\star} \ln \|y_\star(\tau_\star)\| = \lim_{n\to+\infty} \frac{1}{\tau_n} \ln \|y_n(\tau_n)\| = \mathbf{\Lambda}(T,M,k),$$

which completes the proof. ■

Now that we know that the sup in Problem (4.38) is attained, we can use Pontryagin Maximum Principle to characterize the trajectory $y_\star$ that attains the sup of (4.38). We use as reference the statement of Pontryagin Maximum Principle given in Theorem 7.3 of [2], which we recall here.

Theorem 4.10 (Pontryagin Maximum Principle)**.** *Consider the problem*

$$\max_{\gamma\in\mathfrak{U}} \phi_0(\tau, y(\tau)) \tag{4.44a}$$

for the system described by the equations

$$\dot{y} = f(y(t), \gamma(t)), \quad y(0) = y_0, \quad \gamma(t) \in U \text{ almost everywhere,} \tag{4.44b}$$

where the terminal time τ and the terminal point $y(\tau)$ are subject to the constraints

$$\phi_i(\tau, y(\tau)) = 0, \quad i = 1, \ldots, n. \tag{4.44c}$$

We also suppose that f is continuous on $\Omega \times U$, where $\Omega \subset \mathbb{R}^d$ is open, that f is continuously differentiable with respect to y and that the functions ϕ_i, $i = 0, \ldots, n$ are continuously differentiable. Let $\tau_\star$ and $\gamma_\star : [0, \tau_\star] \to U$ maximize (4.44a)*, with $\gamma_\star$ bounded, and let $y_\star$ be the corresponding trajectory, solution of* (4.44b)*. We suppose that the vectors $\left(\frac{\partial \phi_i}{\partial t}, \frac{\partial \phi_i}{\partial y_1}, \ldots, \frac{\partial \phi_i}{\partial y_d}\right)$, $i = 1, \ldots, n$, are linearly independent at the point $(\tau_\star, y_\star(\tau_\star))$. The there exists a nontrivial absolutely continuous row vector p such that*

$$\dot{p}(t) = -p(t) \cdot D_y f(y_\star(t), \gamma_\star(t)), \tag{4.45a}$$

$$p(t) \cdot f(y_\star(t), \gamma_\star(t)) = \max_{\omega \in U} \{p(t) \cdot f(y_\star(t), \omega)\} \tag{4.45b}$$

at almost every time $t \in [0, \tau_\star]$. Moreover, there exist constants $\lambda_0, \ldots, \lambda_n$ with $\lambda_0 \geq 0$ such that

$$p(\tau_\star) = \sum_{i=0}^{n} \lambda_i \nabla \phi_i(\tau_\star, y_\star(\tau_\star)) \neq 0, \tag{4.45c}$$

$$\max_{\omega \in U} \{p(\tau_\star) \cdot f(y_\star(\tau_\star), \omega)\} = -\sum_{i=0}^{n} \lambda_i \frac{\partial \phi_i}{\partial t}(\tau_\star, y_\star(\tau_\star)) \tag{4.45d}$$

where $\nabla \phi_i = \left(\frac{\partial \phi_i}{\partial y_1}, \ldots, \frac{\partial \phi_i}{\partial y_d}\right)$. Finally, the function $t \mapsto p(t) \cdot f(y_\star(t), \gamma_\star(t))$ is constant almost everywhere.

We can then apply Theorem 4.10 to Problem (4.38). The function ϕ_0 in the statement of the theorem and the function f defining the system are

$$\phi_0(t, y) = \frac{1}{t} \ln \|y\|, \qquad f(y, \gamma) = \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma & -h \end{pmatrix} y. \tag{4.46a}$$

Still in the notations of Theorem 4.10, we have

$$U = [0, 1], \qquad \mathcal{U} = L^\infty([0, \tau], [0, 1]). \tag{4.46b}$$

The constraint on the final point can be written as $\phi_1(\tau, y(\tau)) = 0$ with

$$\phi_1(t, y) = y_2, \tag{4.46c}$$

and then $n = 1$. We also remark that f, ϕ_0 and ϕ_1 satisfy the regularity hypothesis stated in the theorem. Then, given $\tau_\star$ and $\gamma_\star$ as in the statement of Lemma 4.9 and the corresponding solution $y_\star$, the conclusions of Theorem 4.10 are valid: there exist a vector p and constants λ_0, λ_1 satisfying (4.45). We now want to gather, from these conclusions, properties that will allow us to characterize $\gamma_\star$ and $y_\star$.

Lemma 4.11. *Let $\tau_\star$, $\gamma_\star$ and $y_\star$ be as in the statement of Lemma 4.9. Then $\gamma_\star$ takes its values on $\{0, 1\}$. Moreover, there exist $s_1, s_2 \in (0, \tau_\star)$ with $s_1 \leq s_2$ such that $\gamma_\star(t) = 1$ if $t \in [0, s_1) \cup (s_2, \tau_\star]$ and $\gamma_\star(t) = 0$ if $t \in (s_1, s_2)$. The solution $y_\star$ is included in the quadrant $Q_2 = \{(y_1, y_2) \mid y_1 \leq 0, y_2 \geq 0\}$ during the interval $[0, s_1]$ and in the quadrant $Q_1 = \{(y_1, y_2) \mid y_1 \geq 0, y_2 \geq 0\}$ during $[s_2, \tau_\star]$.*

Proof. First of all, let us explicitly write the conclusions of Theorem 4.10 in the case of (4.46). We note by p the row vector whose existence is given by Theorem 4.10; the equation (4.45a) it satisfies is

$$\dot{p} = -p \begin{pmatrix} 0 & 1 \\ -3k^{3/2}\gamma_\star(t) & -h \end{pmatrix},$$

that is,

$$\begin{cases} \dot{p}_1(t) = 3k^{3/2}\gamma_\star(t)p_2(t), \\ \dot{p}_2(t) = hp_2(t) - p_1(t). \end{cases} \tag{4.47}$$

We have

$$p \cdot f(y_\star, \omega) = p_1 y_{2\star} - 3k^{3/2}\omega p_2 y_{1\star} - hp_2 y_{2\star},$$

and so the maximization condition (4.45b) is

$$\gamma_\star(t)p_2(t)y_{1\star}(t) = \min_{\omega\in[0,1]} \omega p_2(t)y_{1\star}(t). \tag{4.48}$$

We can now show that $\gamma_\star$ takes its values on $\{0,1\}$. We define the switching function Φ by

$$\Phi(t) = p_2(t)y_{1\star}(t)$$

and then, by (4.48), $\gamma_\star$ can be written in function of Φ as

$$\gamma_\star(t) = \begin{cases} 0 & \text{si } \Phi(t) > 0, \\ 1 & \text{si } \Phi(t) < 0. \end{cases} \tag{4.49}$$

We remark that, if $\Phi(t) \neq 0$ almost everywhere on $[0, \tau_\star]$, the function $\gamma_\star$ is defined almost everywhere by (4.49), and, in particular, it takes its values in $\{0,1\}$. We also remark that Φ is absolutely continuous and

$$\dot{\Phi}(t) = hp_2(t)y_{1\star}(t) - p_1(t)y_{1\star}(t) + p_2(t)y_{2\star}(t);$$

$\dot{\Phi}$ is then also absolutely continuous, which shows that Φ is of class $\mathcal{C}^1$.

We now want to show that Φ cannot be zero on an interval. Indeed, we suppose, by contradiction, that Φ is zero on a certain interval, and we fix an open subinterval J where Φ is zero. For every $s \in J$, we have $\Phi(s) = 0$ and $\dot{\Phi}(s) = 0$. We can then write

$$\Phi(t) = \begin{pmatrix} p_1(t) & p_2(t) \end{pmatrix} \begin{pmatrix} 0 \\ y_{1\star}(t) \end{pmatrix}, \qquad \dot{\Phi}(t) = \begin{pmatrix} p_1(t) & p_2(t) \end{pmatrix} \begin{pmatrix} -y_{1\star}(t) \\ y_{2\star}(t) + hy_{1\star}(t) \end{pmatrix}.$$

By Theorem 4.10, p is nontrivial, and the system (4.47) it satisfies is linear, which means that the vector p is never equal to zero. Conditions $\Phi(s) = 0$ and $\dot{\Phi}(s) = 0$ mean that both $(0, y_{1\star}(s))^{\mathrm{T}}$ and $(-y_{1\star}(s), y_{2\star}(s) + hy_{1\star}(s))^{\mathrm{T}}$ are orthogonal to $p(s)^{\mathrm{T}}$; they are then parallel, which means that $y_{1\star}(s) = 0$ for every $s \in J$. But $\dot{y}_{1\star} = y_{2\star}$, which then shows that $y_{2\star}(s) = 0$ for $s \in J$, and then, by uniqueness of the solution, $(y_{1\star}, y_{2\star})^{\mathrm{T}}$ is the solution equal to 0 over all $\mathbb{R}_+$, which is a contradiction. We thus conclude that Φ cannot be zero on an interval, and thus (4.49) defines $\gamma_\star$ almost everywhere. In particular, $\gamma_\star$ takes its values on $\{0,1\}$.

We now want to show the other results on the statement of the lemma. The method we use here is inspired on the one used on Section 7.3 of [1], where a result of the kind was proved in the case of a time minimization problem of a control system linear in the control. We start by defining the matrices

$$F = \begin{pmatrix} 0 & 1 \\ 0 & -h \end{pmatrix}, \qquad G = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

in such a way that the function f of the control system $\dot{y}(t) = f(y(t), \gamma(t))$ can be written as

$$f(y, \gamma) = Fy - 3k^{3/2}\gamma Gy.$$

We remark also that the switching function Φ and its derivative $\dot{\Phi}$ can be written as

$$\Phi(t) = p(t)Gy_\star(t), \qquad \dot{\Phi}(t) = p(t)[G,F]y_\star(t)$$

where $[G,F] = GF - FG$ is the commutator of the matrices G and F. We define the functions

$$\Delta_A(y) = \det(Fy, Gy) = \begin{vmatrix} y_2 & 0 \\ -hy_2 & y_1 \end{vmatrix} = y_1 y_2,$$
$$\Delta_B(y) = \det(Gy, [G,F]y) = \begin{vmatrix} 0 & -y_1 \\ y_1 & hy_1 + y_2 \end{vmatrix} = y_1^2.$$

The set $\Delta_A^{-1}(0)$, corresponding to the axis y_1 and y_2, is the set of points where the vector fields defined by F and G are parallel and the set $\Delta_B^{-1}(0)$, corresponding to the axis y_2, is the set of points where the vector fields defined by G and $[G,F]$ are parallel. In particular, outside $\Delta_A^{-1}(0)$, Fy and Gy are two linearly independent vectors and they thus constitute a basis of $\mathbb{R}^2$; hence there exist scalars $f_S(y)$ and $g_S(y)$ such that $[G,F]y = f_S(y)Fy + g_S(y)Gy$ for every $y \in \mathbb{R}^2 \setminus \Delta_A^{-1}(0)$. We have $\Delta_B(y) = \det(Gy, [G,F]y) = f_S(y)\det(Gy, Fy) = -f_S(y)\Delta_A(y)$, which shows that

$$f_S(y) = -\frac{\Delta_B(y)}{\Delta_A(y)} = -\frac{y_1}{y_2}.$$

We now want to characterize the switches of $\gamma_\star$ when the trajectory is outside $\Delta_A^{-1}(0) \cup \Delta_B^{-1}(0)$, i.e., when the trajectory is not in one of the axis. We take an open time interval J during which $y_\star$ is outside the axes. In particular, $f_S(y_\star(t))$ and $g_S(y_\star(t))$ are defined for every $t \in J$. If $\gamma_\star$ switches at $t_\star \in J$, Equation (4.49) and the continuity of Φ show that $\Phi(t_\star) = 0$. We then have $p(t_\star)Gy_\star(t_\star) = \Phi(t_\star) = 0$ and thus

$$\dot{\Phi}(t_\star) = p(t_\star)[G,F]y_\star(t_\star) = f_S(y_\star(t_\star))p(t_\star)Fy_\star(t_\star). \tag{4.50}$$

Theorem 4.10 shows that $t \mapsto p(t) \cdot f(y_\star(t), \gamma_\star(t))$ is constant almost everywhere, i.e.,

$$t \mapsto p(t)Fy_\star(t) - 3k^{3/2}\gamma_\star(t)p(t)Gy_\star(t) \tag{4.51}$$

is constant almost everywhere; let us note C this constant. The functions $t \mapsto p(t)Fy_\star(t)$ and $t \mapsto p(t)Gy_\star(t)$ are absolutely continuous, which means that the only times when (4.51) may not be equal C is when $\gamma_\star$ is discontinuous, i.e., the switching times. In particular, by taking the limit as t tends to a switching time by points were $\gamma_\star$ is zero, we obtain that $p(t)Fy_\star(t) = C$ at the switching time, and, since we have $p(t)Gy_\star(t) = \Phi(t) = 0$ in such a time, (4.51) is actually constant everywhere. In particular, equations (4.45b) and (4.45d) show that C is equal

$$C = -\lambda_0 \frac{\partial \phi_0}{\partial t}(\tau_\star, y_\star(\tau_\star)) = \frac{\lambda_0}{\tau_\star^2} \ln \|y_\star(\tau_\star)\|$$

and then $C > 0$. Hence $p(t_\star)Fy_\star(t_\star) \geq 0$; but $p(t_\star)Gy_\star(t_\star) = 0$, $Fy_\star(t_\star)$ and $Gy_\star(t_\star)$ are linearly independent and $p(t_\star) \neq 0$, which shows that $p(t_\star)Fy_\star(t_\star) > 0$, and then, by (4.50), $\dot{\Phi}(t_\star)$ and $f_S(y_\star(t_\star))$ have the same signal. The function $f_S(y_\star(t))$ is different from 0 for every $t \in J$, and thus it keeps a constant signal on this interval. If $f_S(y_\star(t)) > 0$, then $\dot{\Phi}(t) > 0$ for every switching time $t \in J$ of $\gamma_\star$; in particular, this means that $\gamma_\star$ can switch only once in J and that, if it switches, it is, by (4.49), necessarily from 1 to 0. Similarly, if $f_S(y_\star(t)) < 0$, then $\gamma_\star$ switches at most once on J, and this switch can only be from 0 to 1.

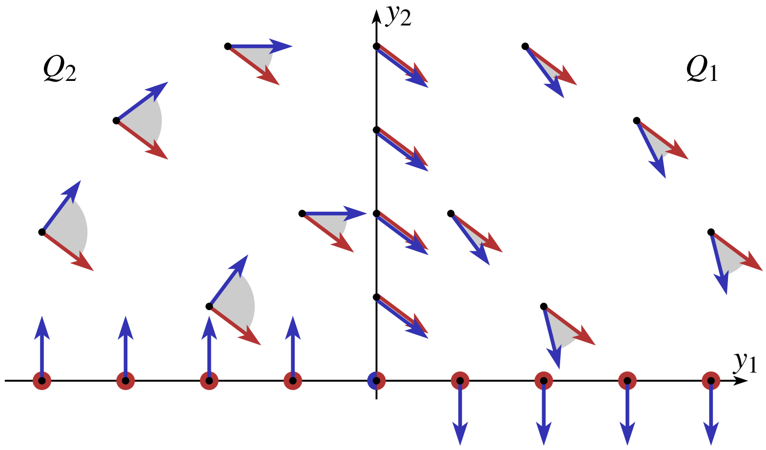

Figure 4.3: Vector field f for $\gamma = 1$ (in blue) and $\gamma = 0$ (in red). All represented vectors are normalized. The conclusions of Pontryagin Maximum Principle imply that $\gamma_\star$ can switch at most once at each the interior of each quadrant Q_1 and Q_2. Moreover, this possible switch is from 1 to 0 in Q_2 and from 0 to 1 in Q_1 since $f_S(y) < 0$ in the interior of Q_1 and $f_S(y) > 0$ in the interior of Q_2.

We can now obtain the desired properties of $\gamma_\star$ and $y_\star$. Starting from $y_\star(0) = (-1, 0)^{\mathrm{T}}$, we can stay stopped in this point if $\gamma_\star(t) = 0$, which cannot maximize ϕ_0, or exit this point going to the interior of Q_2 if $\gamma_\star(0) = 1$; it is thus the second choice that happens, and $\gamma_\star(t) = 1$ in an interval of time around 0. The solution $y_\star$ will eventually exit Q_2, since we have $y_\star(\tau_\star)$ in the nonnegative part of the axis y_1 and this set cannot be reached in finite time from Q_2, and the expressions of the vector fields in the boundaries of Q_2 show that $y_\star$ exits by the axis y_2 and cannot go back inside Q_2; furthermore, it cannot stay stopped in this axis, and thus there exists a unique $s_\star$ such that $y_\star(s_\star)$ is in the axis y_2.

For $t \in (0, s_\star)$, the solution is in the interior of Q_2, where we have $f_S(y) > 0$, and thus we can switch at most once from 1 to 0; we note s_1 the time when this switch happens, with the convention that $s_1 = s_\star$ if the switch does not happen. Now, from $s_\star$ on, the solution goes to the interior of Q_1, until $\tau_\star$ when it reaches the axis y_1, and thus, in the interval $(s_\star, \tau_\star)$, $y_\star$ is in the interior of Q_1, where $f_S(y) < 0$, and thus we can switch at most once from 0 to 1. We remark that, if the solution enters Q_1 with $\gamma_\star(t) = 1$, then it cannot switch anymore, and we will have $\gamma_\star(t) = 1$ until time $\tau_\star$; in this case, we write $s_2 = s_\star$. Otherwise, if the solution enters Q_1 with $\gamma_\star(t) = 0$, it must switch at a certain time $s_2 \in (s_\star, \tau_\star)$ for, if it does not, it cannot reach axis y_1 in a finite time.

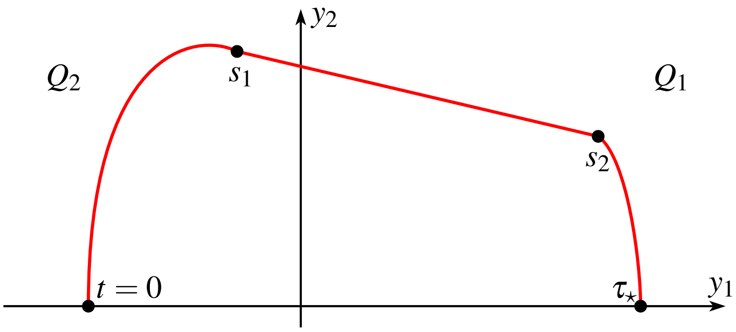

Figure 4.4: Representation of the solution $y_\star$. As stated in Lemma 4.11, $y_\star$ is a solution of (4.41) with $\gamma_\star(t) = 1$ on $[0, s_1)$, $\gamma_\star(t) = 0$ on (s_1, s_2) and $\gamma_\star(t) = 1$ on $(s_2, \tau_\star]$. The solution $y_\star$ lies on Q_2 on $[0, s_1]$ and on Q_1 on $[s_2, \tau_\star]$.

It is now easy to see that, by construction, s_1 and s_2 satisfy the properties stated in the lemma. ■

Now that Lemma 4.11 characterizes $\gamma_\star$ and $y_\star$, Problem (4.38) can be solved more easily. In fact, instead of maximizing the function ϕ_0 of (4.46a) over the whole space of possibles $\tau_\star$ and $\gamma_\star$, we can see that $\gamma_\star$ is completely characterized by the times s_1 and s_2, and thus we have to maximize ϕ_0 over every possible $\tau_\star$, s_1 and s_2, which is a space of dimension 3. The constraint $y_\star(\tau_\star) \in \left\{ \begin{pmatrix} \xi & 0 \end{pmatrix}^{\mathrm{T}}, \, \xi \in \mathbb{R}_+ \right\}$ also reduces the dimension of the space over which we maximize ϕ_0, and then the problem of calculating $\mathbf{\Lambda}(T, M, k)$ reduces to a maximization problem in dimension 2.

Lemma 4.12. *Let $K_5(T, \mu, M)$ be as in Lemma 4.6. There exists $K_6(M)$ such that, if $k > K_5(T, \mu, M)$ and $k > K_6(M)$, then*

$$\mathbf{\Lambda}(T, M, k) \leq \sqrt{3} k^{3/4}. \tag{4.52}$$

Proof. We fix $k > K_5(T, \mu, T)$ and we take $\tau_\star$, $\gamma_\star$ and $y_\star$ as in Lemma 4.9. We then have

$$\mathbf{\Lambda}(T, M, k) = \frac{1}{\tau_\star} \ln \| y_\star(\tau_\star) \| .$$

We use the form of $\gamma_\star$ and $y_\star$ obtained in Lemma 4.11 to estimate this quantity. Let s_1 and s_2 be as in Lemma 4.11. Then, in the interval $[0, s_1]$, we have $\gamma_\star(t) = 1$ and thus $y_\star$ satisfies

$$\dot{y}_\star = \begin{pmatrix} 0 & 1 \\ -3k^{3/2} & -h \end{pmatrix} y_\star, \qquad y_\star(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \tag{4.53}$$

We now take

$$k > \frac{M^2}{36} \tag{4.54}$$

so that $3k^{3/2} > h^2/4$ and we can thus define the positive quantity $\omega = \sqrt{3k^{3/2} - h^2/4}$. A direct calculation shows that the solution of (4.53) is

$$y_{1\star}(t) = -e^{-\frac{h}{2}t} \left(\cos \omega t + \frac{h}{2\omega} \sin \omega t \right), \tag{4.55a}$$

$$y_{2\star}(t) = \left(\omega + \frac{h^2}{4\omega} \right) e^{-\frac{h}{2}t} \sin \omega t. \tag{4.55b}$$

In the interval $[s_1,s_2]$, we have $\gamma_\star(t)=0$ and then the equation satisfied by $y_\star$ is

$$\dot{y}_\star = \begin{pmatrix} 0 & 1 \\ 0 & -h \end{pmatrix} y_\star,$$

which yields the solution

$$y_{1\star}(t) = \frac{1}{h}\left(1-e^{-h(t-s_1)}\right) y_{2\star}(s_2) + y_{1\star}(s_1), \tag{4.56a}$$

$$y_{2\star}(t) = e^{-h(t-s_1)} y_{2\star}(s_1). \tag{4.56b}$$

Finally, in the interval $[s_2,\tau_\star]$, we have $\gamma_\star(t)=1$ and thus the differential equation satisfied by $y_\star$ is the same as in (4.53), but we now consider the condition

$$y_\star(\tau_\star) = \begin{pmatrix} \xi \\ 0 \end{pmatrix}$$

with $\xi>0$. This yields the solution

$$y_{1\star}(t) = \xi e^{-\frac{h}{2}(t-\tau_\star)}\left(\cos\omega(t-\tau_\star) + \frac{h}{2\omega}\sin\omega(t-\tau_\star)\right), \tag{4.57a}$$

$$y_{2\star}(t) = -\xi\left(\omega+\frac{h^2}{4\omega}\right) e^{-\frac{h}{2}t}\sin\omega(t-\tau_\star). \tag{4.57b}$$

We have

$$\mathbf{\Lambda}(T,M,k) = \frac{1}{\tau_\star}\ln\xi. \tag{4.58}$$

To simplify the notation, we write $\sigma=s_2-s_1$. We want to estimate (4.58) in function of s_1 and σ and, to do so, we use the equalities obtained by saying that the solutions given in (4.56) and (4.57) coincide at the point s_2. A development of these equalities yields

$$\xi e^{\frac{h}{2}(\tau_\star-s_2)}\sin\omega(\tau_\star-s_2) = \frac{e^{-h\sigma}}{\omega+\frac{h^2}{4\omega}} y_{2\star}(s_1), \tag{4.59a}$$

$$\xi e^{\frac{h}{2}(\tau_\star-s_2)}\cos\omega(\tau_\star-s_2) = y_{1\star}(s_1) + y_{2\star}(s_1)\left[\frac{1}{h}\left(1-e^{-h\sigma}\right) + \frac{he^{-h\sigma}}{2\omega^2+h^2/4}\right], \tag{4.59b}$$

and we can thus write ξ in function of s_1, σ and $\tau_\star$, and substitution in (4.58) shows that

$$\mathbf{\Lambda}(T,M,k) = \frac{-h(\tau_\star-s_2)+\ln\left[\left(y_{1\star}(s_1)+y_{2\star}(s_1)\left[\frac{1}{h}\left(1-e^{-h\sigma}\right)+\frac{he^{-h\sigma}}{2\omega^2+h^2/4}\right]\right)^2+\left(\frac{e^{-h\sigma}y_{2\star}(s_1)}{\omega+\frac{h^2}{4\omega}}\right)^2\right]}{2[s_1+\sigma+(\tau_\star-s_2)]}. \tag{4.60}$$

To simplify this expression, we first use that $-h(\tau_\star-s_2)\le 0$ and $\tau-s_2\ge 0$. By the expression (4.55b) of $y_{2\star}$ in $[0,s_1]$, we estimate

$$\frac{e^{-h\sigma}y_{2\star}(s_1)}{\omega+\frac{h^2}{4\omega}} \le \sin\omega s_1;$$

we recall that $y_{2\star}(t)\ge 0$, which shows in particular by (4.55b) that $\sin\omega s_1\ge 0$, justifying thus the previous estimate. We also have that $y_{1\star}(s_2)\ge 0$ and $y_{2\star}(s_2)\ge 0$, and then Equations

(4.57) show in particular that $\sin\omega(\tau_\star - s_2) \geq 0$ and $\cos\omega(\tau_\star - s_2) \geq 0$, and then (4.59b) shows that

$$y_{1\star}(s_1) + y_{2\star}(s_1)\left[\frac{1}{h}\left(1 - e^{-h\sigma}\right) + \frac{he^{-h\sigma}}{2\omega^2 + h^2/4}\right] \geq 0;$$

we can then estimate this expression in (4.60) by using that $y_{1\star}(s_1) \leq 0$, which is a conclusion of Lemma 4.11. We also use that $\frac{1}{h}(1 - e^{-h\sigma}) \leq \sigma$ and, by (4.55b), we obtain that

$$y_{2\star}(s_1)\frac{he^{-h\sigma}}{2\omega^2 + h^2/4} \leq \frac{h}{2\omega}\sin\omega s_1.$$

We can estimate $y_{2\star}(s_1)$ by $\left(\omega + \frac{h^2}{4\omega}\right)\sin\omega s_1$ and, by combining all the previous estimates, we obtain that

$$\mathbf{\Lambda}(T,M,k) \leq \frac{\ln(\sin^2\omega s_1) + \ln\left[1 + \left(\sigma\left(\omega + \frac{h^2}{4\omega}\right) + \frac{K}{2\omega}\right)^2\right]}{2(s_1 + \sigma)}.$$

We may also suppose that

$$k > \frac{M^2}{9},$$

which in particular implies (4.54), and, in this case, we have $\frac{h}{2\omega} \leq 1$ and $\omega + \frac{h^2}{4\omega} \leq 2\omega$, which finally yields

$$\mathbf{\Lambda}(T,M,k) \leq \frac{\ln(\sin^2\omega s_1) + \ln\left[1 + (2\omega\sigma + 1)^2\right]}{2(s_1 + \sigma)}.$$

We now define $s' = \omega s_1$, $\sigma' = \omega\sigma$, and then we have

$$\mathbf{\Lambda}(T,M,k) \leq \omega\frac{\ln(\sin^2 s') + \ln\left[1 + (2\sigma' + 1)^2\right]}{2(s' + \sigma')}.$$

A direct calculation shows that the function

$$(s',\sigma') \mapsto \frac{\ln(\sin^2 s') + \ln\left[1 + (2\sigma' + 1)^2\right]}{2(s' + \sigma')}$$

is upper bounded over $(\mathbb{R}_+^*)^2$ and that its upper bound is smaller than 1, and, by bounding ω by $\sqrt{3}k^{3/4}$, we obtain the desired estimate (4.52) under the hypothesis $k > K_5(T,\mu,M)$ and $k > K_6(M)$ with $K_6(M) = M^2/9$. ■

By combining this result with Lemma 4.8, we obtain the desired estimation on the growth of y.

Corollary 4.13. *Let $K_5(T,\mu,M)$ be as in Lemma 4.6 and $K_6(M)$ as in Lemma 4.12. If $k > K_5(T,\mu,M)$ and $k > K_6(M)$, then, for every $\gamma \in \mathfrak{D}(T,\mu,M,k)$ and for every $I = [t_{n-1},t_n] \in \mathfrak{I}_0$, the solution r of* (4.12b) *satisfies*

$$r(t_n) \leq r(t_{n-1})e^{\sqrt{3}k^{3/4}(t_n - t_{n-1})}.$$

4.3.6 Estimate of y

Now that we estimated the growth of y on intervals of the classes $\mathcal{I}_+$ and $\mathcal{I}_0$, we only have to group these results in order to obtain an estimate of the growth of y over an interval $[0,t]$.

Lemma 4.14. *There exists $K_7(T,\mu,M)$ such that, for $k > K_7(T,\mu,M)$, there exists a constant C depending only on T, M and k such that, for every signal $\alpha \in \mathcal{D}(T,\mu,M)$ and every $t \in \mathbb{R}_+$, we can estimate the growth of the solution y of* (4.8) *by*

$$\|y(t)\| \leq C\|y(0)\| e^{2k^{3/4}t}. \tag{4.61}$$

Proof. We suppose that $k > K_i$ for $i = 1,\dots,6$ in order to be in measure to apply all the previous results. Let us fix $\alpha \in \mathcal{D}(T,\mu,M)$ and $t \in \mathbb{R}_+$.

Since the sequence $(t_n)_{n\in\mathbb{N}}$ defined in (4.21) tends monotonically to $+\infty$ as $n \to +\infty$, we see that there exists an $N \in \mathbb{N}$ such that $t \in [t_{N-1},t_N)$ (with the convention $t_{-1} = 0$). We can use Lemma 4.5 and Corollary 4.13 to estimate the growth of y in each interval I_n, $n = 1,\dots,N-1$, but these estimates do not apply to $I_0 = [0,t_0]$ and $[t_{N-1},t]$. The length of these two intervals is however bounded by T, which is a consequence of the proof of Lemma 4.2: we have shown in that lemma that $\theta(t+T)-\theta(t) \leq -2\pi$, and then, in particular, the fact that $\theta(T)-\theta(0) \leq -2\pi$ and the definition of t_0 show that $t_0 \in [0,T)$; for the interval $[t_{N-1},t]$, we see that the fact that $\theta(t_{N-1}+T)-\theta(t_{N-1}) \leq -2\pi$ and the definition of N and t_N show that $t_{N-1} \leq t < t_N < t_{N-1}+T$. We can then use a rough estimation of the growth of y on $[0,t_0]$ and $[t_{N-1},t]$: by Equation (4.12c), we have $\frac{d}{dt}\ln r \leq k^2+h+1$, and then

$$\begin{aligned} r(t_0) &\leq r(0)e^{T(k^2+h+1)}, \\ r(t) &\leq r(t_{N-1})e^{T(k^2+h+1)}. \end{aligned}$$

We now combine these two results with (4.26) and (4.52), which yields

$$\begin{aligned} r(t) \leq e^{2T(k^2+h+1)}r(0) &\left(\prod_{\substack{n=1\\ I_n\in\mathcal{I}_+}}^{N-1} e^{4Mk^{1/2}(t_n-t_{n-1})}\right)\left(\prod_{\substack{n=1\\ I_n\in\mathcal{I}_0}}^{N-1} e^{\sqrt{3}k^{3/4}(t_n-t_{n-1})}\right) \leq \\ &\leq Cr(0)e^{\sqrt{3}k^{3/4}t+4Mk^{1/2}t} \end{aligned}$$

with $C = e^{2T(k^2+h+1)}$, which depends only on T, k and M (by h). It suffices to take k large enough, and more precisely $k \geq \left(\frac{4M}{2-\sqrt{3}}\right)^4$, in order to obtain (4.61). We then take K_7 as the maximum among K_i, $i = 1,\dots,6$, and $\left(\frac{4M}{2-\sqrt{3}}\right)^4$ and the proof is concluded. ■

4.4 Proof of Theorem 4.1

Now that we have studied in details the rate of growth of y, we can prove Theorem 4.1 by combining (4.61) and the relation (4.5) between x and y.

Proof of Theorem 4.1. Let λ be a real constant. We take k satisfying $k > K_7(T,\mu,M)$ and we consider the feedback gain $K = \begin{pmatrix}k^2 & k\end{pmatrix}$. Now, by (4.5), we have that, for every $t \in \mathbb{R}_+$,

$$\|x(t)\| \leq e^{-\frac{k}{2}\int_0^t \alpha(s)ds+\frac{h}{2}t}\left(1+\frac{h}{2}+\frac{k}{2}\right)\|y(t)\|$$

and

$$\|y(t)\| \leq e^{\frac{k}{2}\int_0^t \alpha(s)ds - \frac{h}{2}t}\left(1+\frac{h}{2}+\frac{k}{2}\right)\|x(t)\|,$$

and then, in particular,

$$\|y(0)\| \leq \left(1+\frac{h}{2}+\frac{k}{2}\right)\|x(0)\|.$$

Thus, by combining these with (4.61), we obtain that

$$\|x(t)\| \leq C'\|x(0)\| e^{-\frac{k}{2}\int_0^t \alpha(s)ds + \frac{h}{2}t + 2k^{3/4}t}$$

where C' is a constant depending only on k, M and T. We now use

$$\int_0^t \alpha(s)ds \geq \int_0^{\lfloor \frac{t}{T}\rfloor T} \alpha(s)ds \geq \left\lfloor \frac{t}{T}\right\rfloor \mu \geq \frac{\mu}{T}t - \mu$$

to obtain

$$\|x(t)\| \leq \overline{C}\|x(0)\| e^{(-\frac{k}{2}\frac{\mu}{T}+\frac{h}{2}+2k^{3/4})t}$$

for a new constant $\overline{C}$, which now depends on k, M, T and μ. Now, since, for T, μ and M fixed, we have

$$\lim_{k\to+\infty}\left(-\frac{k}{2}\frac{\mu}{T}+\frac{h}{2}+2k^{3/4}\right) = -\infty,$$

there exists $K(T,\mu,M,\lambda)$ such that, for $k > K(T,\mu,M,\lambda)$, we have $-\frac{k}{2}\frac{\mu}{T}+\frac{h}{2}+2k^{3/4} \leq -\lambda$ and then

$$\|x(t)\| \leq \overline{C}\|x(0)\| e^{-\lambda t}.$$

This concludes the proof, since, for such a k, we have

$$\limsup_{t\to+\infty}\frac{\ln\|x(t)\|}{t} \leq -\lambda.$$

■

We then have a result of convergence at an arbitrary rate for the double integrator when the signal α is a PEL signal.

5 Conclusion

This project helped shedding some light on open problems on switched persistently excited linear systems. The developed result solved Open Problem 5 of [4] for the particular case of the double integrator, using a strategy that consists on treating separately "good" and "bad" intervals of time in order to obtain a fine estimate of the convergence rate of the solution and be able to stabilize it at an arbitrary rate.

A question that arises is whether we can generalize Theorem 4.1. A first possible generalization would be to consider no longer the special case of the double integrator but the general case of a controllable pair $(A,b) \in \mathcal{M}_2(\mathbb{R}) \times \mathbb{R}^2$. We first notice that we only need to consider the case of a matrix A such that $\operatorname{Tr}(A) = 0$, since, for every $\lambda \in \mathbb{R}$, any solution of $\dot{x} = (A - \alpha bK + \lambda \mathrm{Id})x$ can be written on the form $x(t) = e^{\lambda t} y(t)$ where y is a solution of $\dot{y} = (A - \alpha bK)y$, and thus stabilization at an arbitrary rate for one of this systems implies the same result for the other. Up to a linear change of variables, a controllable pair (A,b) with $\operatorname{Tr}(A) = 0$ can be written in the companion form as

$$A = \begin{pmatrix} 0 & 1 \\ -\det(A) & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{5.1}$$

and hence the only necessary generalization to obtain the general case in dimension 2 is to consider a PEL system defined by the matrices (5.1).

We would also like to generalize this result to higher dimensions. The proof that we did here depends in many points on the fact that we are in dimension 2, having no immediate generalization to higher dimensions. In particular, the use of polar coordinates in $\mathbb{R}^2$ can be generalized in dimension d to the use of the coordinates $(r, \omega) \in \mathbb{R}_+^* \times \mathbb{S}^{d-1}$ with $r = \|x\|$ and $\omega = \frac{x}{\|x\|}$; the trajectory of ω is a 1-dimensional immersed submanifold of the $d-1$-dimensional manifold $\mathbb{S}^{d-1}$, and we explored thoroughly in the proof the fact that $1 = d - 1$ in dimension 2, which then gave us many properties of the solution, such as the fact that it rotates around the origin, that the polar angle is monotone in the "good" time intervals and that we can explicitly determine the form of the worst solution in the "bad" time intervals. A generalization using the same technique would need to deal with these problems in higher dimensions.

Regarding the framework of the research project as the final project of the third year of academic studies at École Polytechnique, one can say that the development of the project complements the many courses taken at École Polytechnique. Many results studied during the courses at École Polytechnique were useful in this project, and in particular all the mathematical basis acquired over the years, however the most important aspect of the project was not the direct application of the results of courses, but the fact that it was a practical experience in scientific research in Mathematics. This allowed to see how scientific research is done in practice, starting from the question of understanding the problem and what has already been studied about it before advancing to the part of looking for a solution, which is the challenging part in Mathematics. It is important to know the usual and useful techniques of the domain of the project in order to have ideas of how to look for the solution, but these techniques are not always sufficient and then it is necessary to go deeper and deeper into the problem, understanding its finest details. These details shed more light at the problem and may help to understand why one expects the result to be true, or even give an idea of a counterexample to prove the result false. In any case, creativity, previous experience and the exchange of ideas all play an important role in the search for the solution. In this point of view, the research project complemented the third year courses at École Polytechnique, giving a practical experience on research and developing the required expertise to conclude this academic year.

References

[1] Ugo Boscain et Yacine Chitour: *Introduction à l'automatique*. Notes de cours de MAP561 à l'École Polytechnique, Palaiseau, France, 2011.

[2] Alberto Bressan and Benedetto Piccoli: *Introduction to the Mathematical Theory of Control*. American Institute of Mathematical Sciences, 1st edition, 2007.

[3] Antoine Chaillet, Yacine Chitour, Antonio Loría, and Mario Sigalotti: *Uniform stabilization for linear systems with persistency of excitation: the neutrally stable and the double integrator cases*. Math. Control Signals Systems, 20(2):135–156, 2008, ISSN 0932-4194.

[4] Yacine Chitour and Mario Sigalotti: *On the stabilization of persistently excited linear systems*. SIAM J. Control Optim., 48(6):4032–4055, 2010, ISSN 0363-0129. `http://dx.doi.org/10.1137/080737812`.

[5] Philip Hartman: *Ordinary Differential Equations*. Society for Industrial and Applied Mathematics (SIAM), 2nd edition, 2002.

[6] Daniel Liberzon: *Switching in Systems and Control*. Birkhäuser Boston, 1st edition, 2003.

[7] Michael Margaliot: *Stability analysis of switched systems using variational principles: an introduction*. Automatica, 42(12):2059–2077, 2006.

[8] A. P. Morgan and K. S. Narendra: *On the stability of nonautonomous differential equations $\dot{x} = [A + B(t)]x$ with skew-symmetric matrix $B(t)$*. SIAM J. Control Optim., 15(1):163–176, 1977, ISSN 0363-0129.

[9] H. L. Royden and P. M. Fitzpatrick: *Real Analysis*. Prentice Hall, 4th edition, 2010.

[10] Wilson J. Rugh: *Linear System Theory*. Prentice Hall, 2nd edition, 1996.

[11] Robert Shorten, Fabian Wirth, Oliver Mason, Kai Wulff, and Christopher King: *Stability criteria for switched and hybrid systems*. SIAM Rev., 49(4):545–592, 2007.