



HAL
open science

Rotor position estimation for permanent magnet synchronous electrical machines with salient poles

Jérémy Malaizé, Laurent Praly

► **To cite this version:**

Jérémy Malaizé, Laurent Praly. Rotor position estimation for permanent magnet synchronous electrical machines with salient poles. 2012. hal-02568844v2

HAL Id: hal-02568844

<https://hal.science/hal-02568844v2>

Preprint submitted on 30 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rotor position estimation for permanent magnet synchronous electrical machines with salient poles

Jérémy Malaizé* and Laurent Praly†

written in 2012, revised in april, 2020

Abstract

We extend the one-sided gradient observer of the rotor electrical position of [8] to deal with permanent magnet synchronous electrical machines with salient poles. This is done by an appropriate choice of the output function and involves a Cartesian description of the image of the unit circle by a degree 2 polynomial function and an extension of a convex function. Global convergence of the observer is established under the assumption of limited saliency or small currents.

*J. Malaizé was with the Control and Systems Department, IFP New Energy, Rueil-Malmaison, FRANCE

†L. Praly is with the Centre d'Automatique et Systèmes, ParisTech, École des Mines, Fontainebleau, FRANCE.
laurent.praly@mines-paristech.fr

Foreword

This note reports on a work done by J er emy Malaiz e and myself in 2012. At that time, we were both in the Paris area, J er emy at IFP New Energy in Rueil-Malmaison and myself at CAS, Mines-ParisTech in Fontainebleau. Unfortunately J er emy left for Switzerland. This halted our collaboration being both busy with other things.

The work remained as it was until, getting closer to retirement, I realized it would be a good idea to make it public. Besides helping in the fight against covid 19, the containment has had the other benefit of giving me time to dedicate to this objective.

What the reader will find in the lines below is an edited version of what we wrote in 2012. The modifications concern only the technicalities. All the other aspects – introduction, references, comments, . . . – have not been up-dated although progress has been made on the topic. And worse, the data concerning the successful experiments made at that time on a test bed have been lost and therefore cannot be reported. I apologize for this.

This manuscript has not been submitted for a journal publication and has not been reviewed in any way.

Fontainebleau, April 2020
Laurent Praly

1 Introduction

1.1 The context

The problem of estimating the mechanical state components, i.e. rotor position and speed, for synchronous electrical machines from the measured electrical variables has a very long history [13], rich in various and efficient methods. There are many technical reasons for solving this problem such as cost reduction, wires removal and reliability improvement.

No matter what technique is considered for the control design for synchronous electrical machines, it always comes down to feed an estimate of the rotor position to a state feedback, in charge of controlling the torque delivered by the machine. It is now well established that to design such a feedback, it is easier to work with a two-phase model expressed in a frame rotating with the motor. The so called dq model, that is central to the well-known field-oriented control scheme, see [1] and [13]. However, it turns out that, for the state estimation problem, a model expressed in a fixed frame is more appropriate. This remark was made in [10]. Actually, in that paper it was shown that, by immersing the standard four dimensional model leaving in $\mathbb{R}^3 \times \mathbb{S}^1$ into a five dimensional one leaving in \mathbb{R}^5 , it is possible to describe the dynamics of Surface-Mount Permanent Magnet Synchronous Machines (SM-PMSM) by a system with a triangular structure. In particular, this makes possible the reconstruction of the rotor position from a simple two-dimensional subsystem, completely decoupled from the mechanical behavior of the motor. To be more specific, position estimation may be carried out without any precise knowledge of the mechanical load connected to the machine shaft. There is no need for inertia, friction or load torque knowledge. This is quite an appealing characteristics, since it is usually rather complicated to accurately access these data, subject to changes when the motor is in operation.

This two-dimensional subsystem has been used as a design tool for the gradient observer proposed and studied experimentally in [6], [2] and theoretically in [9]. Unfortunately the study of the corresponding error system, closely related to an averaged approximation of the periodically forced van der Pol oscillator, is very difficult. Also in this literature, only SM-PMSM are considered, and no generalization to other types of synchronous machines is proposed.

In the following, we suggest to modify the gradient observer mentioned above by taking advantage of convexity. The importance of geodesic convexity, with respect to a metric adapted to the system observability, of the level sets of the output function has been pointed out in [11]. Actually convexity also has a long history in estimation and particularly in adaptive control. See [3, 5] for instance. But, in trying to exploit it here, we end up facing a difficulty very similar in spirit to the one presented on [4, page 418]. As in that paper we round it by a conditional correction term that makes the study of the error system much easier than in [9]. Aside from a more accessible convergence proof of the observer, we also propose a generalization of the triangular system structure for SM-PMSM, to a broader class of synchronous machines including for instance Salient-Pole Permanent Magnet Synchronous Machines or Brushless Direct Current Motor exhibiting a trapezoidal back-emf. By combining convexity and this generalization, we provide a general framework for the rotor position estimation for electrical synchronous machines. This approach enjoys some remarkable properties, such as some theoretical guarantees regarding the convergence of the estimated position with robustness to the mechanical load and its characteristics.

The paper is organized as follows. The end of Section 1 introduces the dynamical model considered hereafter, and addresses what kind of electrical machines this approach may be applied to. Section 2 is devoted to the design of an observer for a dynamical system, with an elementary drift and a convex output function. Under some mild assumptions, the proposed observer is proved to yield global asymptotic convergence of the estimated state to the actual system state. In Section 3, we address the problem of estimating the rotor position of a salient-pole permanent magnet synchronous machines. It is showed to be a particular case of the general framework given in Section 2. Finally in Section 4 we show how to extend the previous results to electrical machines with a back-emf assumed

only to be periodic in the rotor position.

1.2 System modeling and problem statement

In the following, the equations of AC synchronous machines cast in the so-called stationary frame are considered. By making use of the Faraday's and Joule's Laws, the phase-to-neutral voltages at the motor terminals read [1, 14] :

$$v = Ri + \dot{\varphi} \quad (1)$$

This allows us to relate the voltages v at the motor terminals to the derivative of the total flux φ encompassed by the windings, and to the currents i within them, while R stands for the stator windings resistance. For three-phase motors, without any assumption regarding the kind of machines considered, the quantities v , i as well as φ are two dimensional vectors obtained *via* the Concordia Transform applied to the three phase-to-neutral voltages, the three line currents and the total flux seen by each of the three phases, respectively.

Actually, in AC synchronous machines the total flux must satisfy (1) in such a way that it remains a prescribed function of the currents i as well as of the rotor electrical phase, denoted θ . Here we shall investigate more the case of salient-pole synchronous machines for which the total flux satisfies[14] :

$$\varphi = L_0 i + \Phi \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + L_1 \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} i, \quad (2)$$

where L_0 and L_1 are some inductances, while Φ is homogeneous to a magnetic flux :

- When L_1 is zero, (2) degenerates to the modeling of surface-mount synchronous machines, addressed in [8],
- When Φ equals zero, (2) describes the behavior of synchro-reluctant machines,
- For permanent magnet synchronous machine featuring saliency, Φ is considered to be constant, up to the effects of the temperature changes,
- For wound-rotor synchronous machine, Φ is a time-varying control input.

Our problem is the determination of the electrical phase θ from the only knowledge of v and i , and under a perfect knowledge of the motor physical parameters, namely R , Φ , L_0 and L_1 . The usual solution would go with first rewriting (1) and (2) as :

$$\dot{x} = f(x, (v, \omega)) \quad , \quad y = i$$

with the state x made of i and θ , the measured output y made of i and the input made of v and ω , the rotor speed. But this has the drawback of asking for either the knowledge of ω or for a model generating it, with involving possibly the mechanics of the motor and the knowledge of the motor load, friction or inertia. Instead we follow the same approach as the one proposed in [10] for surface-mount PMSM. For such motors, L_1 equals zero in (2), and by eliminating θ , we find that the total flux is constrained to satisfy at each time

$$\mathcal{C}(\varphi, t) = 0 \quad , \quad (3)$$

where

$$\mathcal{C}(\varphi, t) = |\varphi - L_0 i(t)|^2 - \Phi^2 \quad .$$

This says the total flux evolves on a circle with radius Φ and moving center $L_0 i$. This motivates for considering the dynamical system

$$\dot{\varphi} = v - Ri \quad , \quad y = \mathcal{C}(\varphi, t) \quad (4)$$

with state φ , input $v - Ri$ and measurement y which turns out to be identically zero for signals coming from the motor. Below we go exactly the same way. We prove that, in the case of (2) or even more generally, in the case where the total flux is expressed as the following truncated Fourier series expansion of order N in θ , i.e.

$$\varphi = \sum_{n=0}^N \varphi_n(t) \begin{pmatrix} \cos(n\theta + \psi_n(t)) \\ \sin(n\theta + \psi_n(t)) \end{pmatrix}, \quad (5)$$

where, for any n between 0 and N , $\varphi_n : \mathbb{R}_{\geq 0} \mapsto \mathbb{R}$ and $\psi_n : \mathbb{R}_{\geq 0} \mapsto [0, 2\pi[$ define the magnitude and the phase of the n -th term of the expansion, respectively, the constraint function \mathcal{C} is a polynomial in φ , with higher degree term $|\varphi|^{2N}$, and the coefficients of which depend on the $\varphi_n(t)$'s (through polynomials) and the $\psi_n(t)$'s (through sine and cosine functions).

With this, it remains to build an observer for the system (4). This is what we do in the next section.

2 An observer for a system with an elementary drift and a convex output function

This section is nothing but a reproduction of [8, Section II]. Its object is the proposition and the analysis of an observer for a generic simple system. It will be the basis for the observer of the total flux, we present in the next section.

2.1 Observer design

Consider a dynamical system with state x of dimension two, the dynamics of which simply reads:

$$\dot{x} = u(t), \quad (6)$$

with x in \mathbb{R}^2 , $u : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^2$ a continuous function the value $u(t)$ of which is known at each time t . Assume further that u as well as its time integral are bounded on $\mathbb{R}_{\geq 0}$. Denote by $t \mapsto \mathcal{X}(x_0, t; u)$ the solution of this dynamical system issued from x_0 at time 0, under the action of u . According to the assumptions, this solution is defined and bounded on \mathbb{R}_+ .

The information about x is given by the measurement :

$$y(t) = h(x, t),$$

where $h : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$ is an output function. Actually we restrict our attention to what we call constrained solutions. A solution $t \mapsto \mathcal{X}_c(x_0, t; u)$ is called a constrained solution if its measurement y is zero for all time, i.e. we have :

$$h(\mathcal{X}_c(x_0, t; u), t) = 0 \quad \forall t \in \mathbb{R}_{\geq 0}. \quad (7)$$

or equivalently :

$$h(x_0, 0) = 0, \quad \frac{\partial h}{\partial x}(\mathcal{X}_c(x_0, t; u), t)u(t) + \frac{\partial h}{\partial t}(\mathcal{X}_c(x_0, t; u), t) = 0 \quad \forall t \in \mathbb{R}_{\geq 0}. \quad (8)$$

We assume :

A1: The function $h : \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ is continuous, the function $x \mapsto h(x, t)$ is C^2 for each $(x, t) \mapsto h(x, t)$

t and, for any positive scalar c , there exist positive real numbers $\bar{h}(c)$, $\bar{h}_x(c)$ and $\bar{h}_t(c)$ such that :

$$|h(x, t)| \leq \bar{h}(c), \quad \left| \frac{\partial h}{\partial x}(x, t) \right| \leq \bar{h}_x(c), \quad \left| \frac{\partial h}{\partial t}(x, t) \right| \leq \bar{h}_t(c) \quad \forall (x, t) : |x| \leq c, \quad t \geq 0.$$

A2: There exists a strictly positive scalar \underline{h}_{xx} such that the Hessian matrix of h satisfies :

$$\frac{\partial^2 h}{\partial x \partial x}(x, t) \geq 2\underline{h}_{xx}I, \quad \forall (x, t) : h(x, t) \geq 0, t \geq 0.$$

A3: For any constrained solution $t \mapsto \mathcal{X}_c(x_0, t; u)$ of (6), and any unit vector w in \mathbb{R}^2 , there exists a sequence t_n going to infinity as n grows indefinitely such that:

$$\frac{\partial h}{\partial x}(\mathcal{X}_c(x_0, t_n; u), t_n) w > 0.$$

Remark 1.

1. Besides some smoothness properties, A1 simply requires that the functions $x \mapsto h(x, t)$ and $x \mapsto \frac{\partial h}{\partial x}(x, t)$ are bounded, uniformly in t , on any compact subset of \mathbb{R}^2 .
2. Under Assumption A2, we have

$$h(x_2, t) \geq h(x_1, t) + \frac{\partial h}{\partial x}(x_1, t)[x_2 - x_1] + \underline{h}_{xx}|x_1 - x_2|^2 \quad \forall (x_1, x_2, t) \quad (9)$$

This is a direct consequence of Taylor's expansion with integral remainder, i.e. :

$$h(x_2, t) = h(x_1, t) + \frac{\partial h}{\partial x}(x_1, t)[x_2 - x_1] + \frac{1}{2}[x_2 - x_1]^\top \left[\int_0^1 \int_0^1 \left(\frac{\partial^2 h}{\partial x^2}(x_2 + rs(x_2 - x_1), t) dr \right) s ds \right] [x_2 - x_1],$$

with using the lower bound for $\frac{\partial^2 h}{\partial x^2}(x_2 + rs(x_2 - x_1), t)$ given by A2. Hence Assumption A2 requires the function $x \mapsto h(x, t)$ is strongly convex, uniformly in t and therefore the sub-level set $\{x : h(x, t) \leq 0\}$ is strictly convex for each $t \geq 0$.

3. If A3 is fulfilled, the gradient of h evaluated along any constrained solution $t \mapsto \mathcal{X}_c(x_0, t; u)$ of (6) points infinitely often in any direction of \mathbb{R}^2 . This is a non-evanescent full spanning condition.

We propose the following observer for the state x :

$$\dot{\hat{x}} = u(t) - \mu \frac{\partial h}{\partial x}(\hat{x}, t)^\top \max\{0, h(\hat{x}, t)\}, \quad (10)$$

where μ is an arbitrary strictly positive real number. It is an algorithm of a gradient type but with a correction term which is “on” only when $h(\hat{x}, t)$ is non negative. In the following, we denote $\hat{\mathcal{X}}(\hat{x}_0, t; u)$ its solution at time t , issued from \hat{x}_0 at time 0. Because of A1 it is well defined and unique. Note that it depends on the actual constrained system solution $\mathcal{X}_c(x_0, t; u)$ indirectly only via u satisfying (8).

2.2 Observer convergence

Proposition 1. *Under assumptions A1, A2 and A3, for any (x_0, \hat{x}_0) , we have :*

$$\left| \mathcal{X}_c(x_0, t; u) - \hat{\mathcal{X}}(\hat{x}_0, t; u) \right| \leq |x_0 - \hat{x}_0| \quad \forall t \geq 0 \quad (11)$$

and :

$$\lim_{t \rightarrow \infty} \left| \mathcal{X}_c(x_0, t; u) - \hat{\mathcal{X}}(\hat{x}_0, t; u) \right| = 0. \quad (12)$$

This Proposition says that the zero error set $\mathfrak{r} = \{(x, \hat{x}) \in \mathbb{R}^2 : x = \hat{x}\}$ is globally and asymptotically stable but only when restricted to constrained solutions.

Proof. The error dynamics is obtained by combining (6) and (10), and simply reads :

$$\overbrace{\hat{x} - x}^{\dot{\phantom{\hat{x} - x}}} = -\mu \frac{\partial h}{\partial x}(\hat{x}, t) \max\{0, h(\hat{x}, t)\}. \quad (13)$$

Let us consider :

$$V = \frac{1}{2}[\hat{x} - x]^\top [\hat{x} - x]$$

as a Lyapunov function candidate. Its time derivative is :

$$\dot{V} = -\mu \max\{0, h(\hat{x}, t)\} \frac{\partial h}{\partial x}(\hat{x}, t) (\hat{x} - x).$$

With (9), this yields :

$$\dot{V} \leq -2\mu \underline{h}_{xx} V \max\{0, h(\hat{x}, t)\} \quad \forall (x, \hat{x}, t) : h(x, t) = 0.$$

This implies (11).

Moreover, the evaluation of V along the solution at time t which we denote $V(t)$ satisfies :

$$V(t) \leq \exp\left(-2\mu \underline{h}_{xx} \int_0^t \max\{0, h(\hat{\mathcal{X}}(\hat{x}_0, s; u), s)\} ds\right) V(0).$$

The limit (12) follows when the following integral diverges :

$$I(t) = \int_0^t \max\{0, h(\hat{\mathcal{X}}(\hat{x}_0, t; u), s)\} ds.$$

Let us establish this holds by contradiction. We assume that $V(t)$ does not converge to 0. This implies that I converges. Since, by assumption, the integral of $t \mapsto u(t)$ is bounded, the function $t \mapsto \mathcal{X}_c(x_0, t; u)$ is bounded. Then (11) implies the existence of a solution dependent positive scalar c such that :

$$|\mathcal{X}_c(x_0, t; u)| \leq c, \quad |\hat{\mathcal{X}}(\hat{x}_0, t; u)| \leq c \quad \forall t \geq 0.$$

On another hand A1 and (13) give :

$$|\overbrace{\hat{x} - x}^{\dot{\phantom{\hat{x} - x}}}| \leq \mu \bar{h}_x(c) \max\{0, h(\hat{x}, t)\} \quad \forall (x, \hat{x}, t) : |\hat{x}| \leq c, t \geq 0.$$

With the notation :

$$\tilde{\mathcal{X}}(t) = \hat{\mathcal{X}}(\hat{x}_0, t; u) - \mathcal{X}_c(x_0, t; u),$$

this yields, for any $0 \leq t_1 \leq t_2$,

$$|\tilde{\mathcal{X}}(t_2) - \tilde{\mathcal{X}}(t_1)| \leq \int_{t_1}^{t_2} |\dot{\tilde{\mathcal{X}}}(s)| ds \leq \mu \bar{h}_x(c) |I(t_2) - I(t_1)|.$$

The convergence of $I(t)$ implies that, for any sequence t_n tending to infinity, $I(t_n)$ is a Cauchy sequence, and so is $\tilde{\mathcal{X}}(t_n)$ as a consequence of the previous inequality. It follows that, as t tends to infinity, $\tilde{\mathcal{X}}(t)$ admits a limit, denoted $\tilde{\mathcal{X}}_*$, which is different from zero by assumption. Rewrite $\tilde{\mathcal{X}}$ as :

$$\tilde{\mathcal{X}}(t) = \tilde{\mathcal{X}}_* + \varepsilon_1(t),$$

with $\varepsilon_1 : \mathbb{R}_+ \rightarrow \mathbb{R}^2$ tending to zero.

According to A1 and the fact that the function $t \mapsto \hat{\mathcal{X}}(\hat{x}_0, t; u)$ is bounded, the function $t \mapsto \hat{\mathcal{X}}(\hat{x}_0, t; u)$ is uniformly continuous, and the function $t \mapsto \max \{0, h(\hat{\mathcal{X}}(\hat{x}_0, t; u), t)\}$ has the same property. For I converges, Barbalat's Lemma may be invoked to come up with :

$$\lim_{t \rightarrow \infty} \varepsilon_2(t) = 0 ,$$

where $\varepsilon_2(t)$ is defined as :

$$\varepsilon_2(t) = \max \left\{ 0, h(\hat{\mathcal{X}}(\hat{x}_0, t; u), t) \right\} \quad \left(\geq h(\hat{\mathcal{X}}(\hat{x}_0, t; u), t) \right) .$$

From the definition of $\tilde{\mathcal{X}}_*$ and ε_1 , this translates into :

$$\varepsilon_2(t) \geq h(\mathcal{X}_c(x_0, t; u) + \tilde{\mathcal{X}}_* + \varepsilon_1(t), t) .$$

Now (9) together with :

$$h(\mathcal{X}_c(x_0, t; u), t) = 0 \quad \forall t \geq 0$$

give :

$$\varepsilon_2(t) \geq h(\mathcal{X}_c(x_0, t; u) + \tilde{\mathcal{X}}_* + \varepsilon_1(t), t) \geq \frac{\partial h}{\partial x}(\mathcal{X}_c(x_0, t; u), t) \left(\tilde{\mathcal{X}}_* + \varepsilon_1(t) \right) + \underline{h}_{xx} |\tilde{\mathcal{X}}_* + \varepsilon_1(t)|^2 .$$

Since h satisfies A1, $t \mapsto \mathcal{X}_c(x_0, t; u)$ is bounded, $\tilde{\mathcal{X}}_*$ is non-zero and ε_1 and ε_2 tend to zero. There exists a time $T > 0$ such that we have :

$$-\frac{1}{2} \underline{h}_{xx} |\tilde{\mathcal{X}}_*|^2 \geq \frac{\partial h}{\partial x}(\mathcal{X}_c(x_0, t; u), t) \tilde{\mathcal{X}}_* \quad \forall t \geq T .$$

This is in contradiction with A3. So we must have :

$$\lim_{t \rightarrow \infty} V(t) = 0 .$$

This implies (12). □

3 Rotor position estimation for permanent magnet synchronous electrical machines with salient poles

From the design given in the previous section, a total flux observer for a permanent magnet synchronous electrical machine with salient poles is :

$$\dot{\hat{\varphi}} = v - Ri + \mu \frac{\partial h}{\partial \varphi}(\hat{\varphi}, t)^\top \max \{0, h(\hat{\varphi}, t)\} , \quad (14)$$

provided we can choose the output function h such that conditions A1 to A3 hold and the equation :

$$h(\varphi, t) = 0$$

is satisfied by the total flux given by any solution to (1) which is constrained, i.e. there exists θ such that (2) holds. In Proposition 3 of Section 3.2 below, we establish such a choice is possible and leads to the observer :

$$\dot{\hat{\varphi}} = v - Ri \quad \text{if} \quad \left[\left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi^2 |\hat{\varphi} - (L_0 - L_1) i(t)|^2 \right] \leq 0$$

and :

$$\begin{aligned} \dot{\hat{\varphi}} = & v - Ri \\ & + \mu \left[4 \left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right) (\hat{\varphi} - L_0 i(t)) - 2\Phi^2 (\hat{\varphi} - (L_0 - L_1)i(t)) \right] \times \\ & \times \left[\left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi^2 |\hat{\varphi} - (L_0 - L_1)i(t)|^2 \right] \end{aligned}$$

if not. Its global convergence will be established in Proposition 3, under the assumption that the electrical angular position does not stop. But before this, we describe how the particular output function h is obtained.

3.1 Image of the unit circle by a degree 2 polynomial function

We have :

Proposition 2. *If φ is given by (2), then, for each t , it satisfies*

$$\mathcal{C}(\varphi, t) = 0$$

where :

$$\mathcal{C}(\varphi, t) = \left(|\varphi - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi |\varphi - (L_0 - L_1)i(t)|^2 . \quad (15)$$

Moreover, when :

$$\Phi > 2L_1|i(t)| . \quad (16)$$

we cannot have :

$$\varphi = (L_0 - L_1)i(t) .$$

Conversely, assume (16) holds. If φ satisfies :

$$\mathcal{C}(\varphi, t) = 0 \quad \text{and} \quad \varphi \neq (L_0 - L_1)i(t) , \quad (17)$$

then φ satisfies (2), with θ given by :

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \frac{\varphi - (L_0 - L_1)i(t)}{|\varphi - (L_0 - L_1)i(t)|} . \quad (18)$$

Proof. To ease the proof, we use the complex representation, with symbols in bold style type, of vectors in \mathbb{R}^2 and introduce the following notations :

- $\boldsymbol{\varphi} = \varphi_\alpha + \boldsymbol{j}\varphi_\beta \in \mathbb{C}$ is the complex representation of the total flux vector $\varphi = (\varphi_\alpha \ \varphi_\beta)^\top$,
- $\boldsymbol{i} = i_\alpha + \boldsymbol{j}i_\beta \in \mathbb{C}$ is the complex representation of the current vector $i = (i_\alpha \ i_\beta)^\top$,
- $\boldsymbol{\xi} = \cos(\theta) + \boldsymbol{j}\sin(\theta)$,
- $\boldsymbol{\gamma}_0 = L_0 \boldsymbol{i}$, $\boldsymbol{\gamma}_1 = \Phi$, $\boldsymbol{\gamma}_2 = L_1 \bar{\boldsymbol{i}}$

where $\bar{\cdot}$ is the conjugate complex.

With these notations, the fact that φ satisfies (2) is equivalent to the fact that the polynomial :

$$P(\xi) = (\gamma_0 - \varphi) + \gamma_1 \xi + \gamma_2 \xi^2$$

has a root on the unit circle \mathbb{S}^1 .

Necessary condition :

To obtain a necessary condition, we note that the equation :

$$\varphi(\xi) = \gamma_0 + \gamma_1 \xi + \gamma_2 \xi^2$$

defines a function $\varphi : \mathbb{S}^1 \mapsto \mathbb{C}$, the image of which is a closed planar curve. Actually ξ is a parameter giving a parametric expression of this curve. Let us derive its Cartesian equation. For this, we eliminate ξ from the relationship between the real φ_α and imaginary φ_β parts of φ .

Since ξ is on the unit circle, we have

$$\bar{\xi}\xi = 1 .$$

Then, by multiplying P by $\bar{\xi}^2$ and taking the conjugate complex of the resulting expression leads to the other polynomial :

$$Q(\xi) = \bar{\gamma}_2 + \bar{\gamma}_1 \xi + (\bar{\gamma}_0 - \bar{\varphi}) \xi^2 .$$

By construction, if ξ is a root of P , it is a root of Q . So these two polynomials have a common root, which is equivalent to the property that their resultant is zero, see [12, chapter 4]. Here, this resultant is the determinant of the following block-matrix :

$$M = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \bar{\mathbf{B}}^\top & \bar{\mathbf{A}}^\top \end{pmatrix} .$$

where \mathbf{A} and \mathbf{B} are the following two Toeplitz matrices

$$\mathbf{A} = \begin{pmatrix} \gamma_0 - \varphi & \gamma_1 \\ 0 & \gamma_0 - \varphi \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \gamma_2 & 0 \\ \gamma_1 & \gamma_2 \end{pmatrix} .$$

They satisfy :

$$\mathbf{A}\bar{\mathbf{B}}^\top = \bar{\mathbf{B}}^\top \mathbf{A}$$

and therefore :

$$\begin{pmatrix} \mathbf{I} & 0 \\ -\bar{\mathbf{B}}^\top & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \bar{\mathbf{B}}^\top & \bar{\mathbf{A}}^\top \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{A}\bar{\mathbf{A}}^\top - \bar{\mathbf{B}}^\top \mathbf{B} \end{pmatrix} .$$

We obtain algebraically :

$$\det(M) = \det(\mathbf{A}\bar{\mathbf{A}}^\top - \bar{\mathbf{B}}^\top \mathbf{B}) = [|\gamma_0 - \varphi|^2 - |\gamma_2|^2]^2 - |\bar{\gamma}_1[\gamma_0 - \varphi] - \bar{\gamma}_2\gamma_1|^2 . \quad (19)$$

With the expressions above of the γ_i , (15) is nothing but the equation $\det(M) = 0$.

When $\gamma_1 \neq 0$, a candidate solution to this equation could be :

$$\varphi = \gamma_0 - \bar{\gamma}_2 \frac{\gamma_1}{\bar{\gamma}_1}$$

Indeed in this case, we get the 2 equations :

$$0 = |\gamma_0 - \varphi|^2 - |\gamma_2|^2, \quad 0 = \bar{\gamma}_1[\gamma_0 - \varphi] - \bar{\gamma}_2\gamma_1 .$$

But then, if ξ is the root of P with modulus 1, we get successively :

$$\begin{aligned} 0 &= \bar{\gamma}_1(\gamma_0 - \varphi) + |\gamma_1|^2 \xi + \bar{\gamma}_1 \gamma_2 \xi^2, \\ &= \bar{\gamma}_2 \gamma_1 + |\gamma_1|^2 \xi + \bar{\gamma}_1 \gamma_2 \xi^2, \\ &= \bar{\gamma}_2 \gamma_1 \bar{\xi} + |\gamma_1|^2 + \bar{\gamma}_1 \gamma_2 \xi \end{aligned}$$

and therefore :

$$|\gamma_1|^2 = -2\Re(\bar{\gamma}_2 \gamma_1 \bar{\xi}) \leq 2|\gamma_2||\gamma_1|.$$

So the candidate solution is not a solution when :

$$|\gamma_1| > 2|\gamma_2|.$$

Sufficient condition :

Let ρ and θ be the polar coordinates of the vector $\varphi - (L_0 - L_1)i(t)$ assumed not zero, i.e. :

$$\varphi - (L_0 - L_1)i(t) = \rho \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

(17) gives :

$$\begin{aligned} \left(|\varphi - (L_0 - L_1)i(t)|^2 - 2L_1 i(t)^\top [\varphi - (L_0 - L_1)i(t)] \right)^2 &= \Phi^2 |\varphi - (L_0 - L_1)i(t)|^2, \\ \left(\rho^2 - 2\rho L_1 i(t)^\top \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \right)^2 &= \Phi^2 \rho^2, \\ \left(\rho - 2L_1 i(t)^\top \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \right)^2 &= \Phi^2. \end{aligned}$$

So ρ is :

$$(0 \leq) \quad \rho = 2L_1 i(t)^\top \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \pm \Phi.$$

The minus case is impossible because of (16). So we have :

$$\varphi - (L_0 - L_1)i(t) = \left[2L_1 i(t)^\top \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + \Phi \right] \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

And, with the identity :

$$-i + 2 \begin{pmatrix} \cos(\theta) & \sin(\theta) \end{pmatrix} i \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} i,$$

we have obtained (2). □

3.2 Choice of the output function h

Proposition 2 says the total flux φ given, at time t , by any constrained solution is in the set $\{\varphi : \mathcal{C}(\varphi, t) = 0\}$ with the constraint function \mathcal{C} given as :

$$\mathcal{C}(\varphi, t) = \left(|\varphi - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi^2 |\varphi - (L_0 - L_1)i(t)|^2 \quad \forall(\varphi, t).$$

It says also the total flux :

$$\varphi = (L_0 - L_1)i(t),$$

although in this set, should not be considered if :

$$\Phi > 2L_1|i(t)| .$$

But, importantly, it says nothing about pairs (φ, t) which are not in the set $\{(\varphi, t) : \mathcal{C}(\varphi, t) = 0\}$. This fact is exploited below in (30).

From this information we are tempted to choose the output function h to be used in the observer satisfying :

$$\left\{ \varphi : h(\varphi, t) = 0 \right\} = \left\{ \varphi : \mathcal{C}(\varphi, t) = 0 , \quad \varphi \neq (L_0 - L_1)i(t) \right\} \quad \forall t : \Phi > 2L_1|i(t)| .$$

But how to choose h outside the set $\{\varphi : \mathcal{C}(\varphi, t) = 0\}$? Our answer is

Proposition 3. *Assume there exists a strictly positive real number φ in $\left]0, \frac{1}{2}\right]$ such that we have :*

$$2\frac{L_1}{\Phi}|i(t)| \leq \frac{1}{2} - \varphi \quad \forall t > 0 .$$

Then, with the notation :

$$\mathcal{C}(\varphi, t) = \left(|\varphi - L_0i(t)|^2 - |L_1i(t)|^2 \right)^2 - \Phi^2 |\varphi - (L_0 - L_1)i(t)|^2 ,$$

there exists a function h satisfying Conditions A1 and A2 and¹ :

$$h(\varphi, t) = \mathcal{C}(\varphi, t) \quad \forall (\varphi, t) : \mathcal{C}(\varphi, t) \geq 0 \quad \text{and} \quad |\varphi - L_0i(t)|^2 - L_1^2|i(t)|^2 \geq \frac{3\Phi^2 - 4\Phi L_1|i(t)|}{4} , \quad (20)$$

To prove this result, we start by summarizing here the results about the constraint function \mathcal{C} which are established in appendix B. This is made easier by changing the coordinate φ of the total flux in :

$$z = \frac{\varphi}{\Phi} - \frac{L_0}{\Phi}i(t) + \frac{L_1}{\Phi}i(t) \quad (21)$$

and by introducing the notation :

$$e(t) = 2\frac{L_1}{\Phi}i(t) .$$

The expression of the constraint function \mathcal{C} with this other coordinate is :

$$\Phi^2 \mathcal{C}(\varphi, t) = \mathcal{C}_z(z, t) = \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 . \quad (22)$$

It is also appropriate to introduce the e -dependent function $\bar{\eta} : \left[\frac{3-2|e(t)|}{4}, 1 - |e(t)| \right] \mapsto \mathbb{R}$ as :

$$\bar{\eta}(\bar{\mathbf{h}}) = \bar{\mathbf{h}}^2 - \frac{1}{4} \left[|e(t)| - \sqrt{|e(t)|^2 + 4\bar{\mathbf{h}}^2} \right]^2 . \quad (23)$$

Its properties, established in Appendix A, are :

¹The last inequality is : $|z|^2 - e(t)^\top z \geq \frac{3-2|e(t)|}{4}$.

Lemma 1. *The function $\bar{\eta}$ takes strictly negative values on the interval $\left[\frac{3-2|e(t)|}{4}, 1-|e(t)|\right]$, is strictly increasing on the open interval $\left[\frac{4-e^2-e\sqrt{e^2+8}}{8}, +\infty\right]$ and satisfies :*

$$\bar{\eta}\left(\frac{3-2|e(t)|}{4}\right) \leq -\frac{\wp}{12} \quad , \quad \bar{\eta}(\bar{h}_a) - \bar{\eta}(\bar{h}_b) \geq \frac{\bar{h}_a - \bar{h}_b}{3} > 0 \quad \forall (\bar{h}_a, \bar{h}_b) : \bar{h}_a \geq \bar{h}_b \geq \frac{3-2|e(t)|}{4} .$$

Finally we denote \mathcal{D} the set :

$$\mathcal{D} = \left\{ (z, t) \in \mathbb{R}^2 \times \mathbb{R}_{\geq 0} : |z|^2 - e(t)^\top z > \frac{3-2|e(t)|}{4} \right\} .$$

For t fixed, this is \mathbb{R}^2 from which we have excluded :

$$\mathcal{S}_{ex} = \left\{ (z, t) : |z|^2 - e(t)^\top z \leq \frac{3-2|e(t)|}{4} \right\} ,$$

which is the disk with center $\frac{e(t)}{2}$ and radius $\sqrt{\frac{3-2|e(t)|+|e(t)|^2}{4}}$.

Proposition 4. *Let \mathcal{C}_z be the constraint function defined in (22), i.e. :*

$$\mathcal{C}_z(z, t) = \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2$$

1. *The equation $\mathcal{C}_z(z, t) = 0$ is the Cartesian form of a Limaçon of Pascal. See [7, pp. 44-51]. For $|e(t)| \leq \frac{1}{2}$, it encircles a convex set. With point 2 in Remark 1, this motivates :*

Assumption H: There exists a strictly positive real number \wp in $\left]0, \frac{1}{2}\right]$ such that we have² :

$$2\frac{L_1}{\Phi}|i(t)| = |e(t)| \leq \frac{1}{2} - \wp \quad \forall t > 0 .$$

2. *Under Assumption H, we have :*

$$|z_a| \leq \frac{3}{2} \quad , \quad |z_a - z_b| \geq \frac{1}{15} |\mathcal{C}_z(z_a, t) - \mathcal{C}_z(z_b, t)|$$

for all (z_a, z_b, t) satisfying :

$$\mathcal{C}_z(z_a, t) \leq 0 \quad , \quad \mathcal{C}_z(z_b, t) \leq 0 .$$

3. *Under Assumption H, for all \bar{h} in $\left[\frac{3-2|e(t)|}{4}, 1-|e(t)|\right]$, all (z, t) in the super level set :*

$$\{(z, t) \in \mathcal{D} : \mathcal{C}_z(z, t) \geq \bar{\eta}(\bar{h})\} ,$$

satisfy :

$$\frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z, t) \geq 2\wp I \quad , \quad d\left(z, \left\{ \zeta \in \mathbb{R}^2 : |\zeta|^2 - e(t)^\top \zeta \leq \frac{3-2|e(t)|}{4} \right\}\right) \geq \frac{\bar{h} - \frac{3-2|e(t)|}{4}}{2} . \quad (24)$$

with $d(\cdot, \cdot)$ denoting the Euclidean distance.

²This says the saliency or the current should be small.

4. Under Assumption H, the function \mathcal{C}_z is globally convex on the super level set :

$$\left\{ (z, t) \in \mathcal{D} : \mathcal{C}_z(z, t) \geq \bar{\eta} \left(\frac{3 - 2|e(t)|}{4} \right) \right\} .$$

Namely for each t , each z_a and z_b such that (z_a, t) and (z_b, t) are in the super level set above and for all s in $[0, 1]$ such that, with :

$$z_s = sz_a + (1 - s)z_b,$$

(z_s, t) is also in the super level set, we have :

$$s\mathcal{C}_z(z_a, t) + (1 - s)\mathcal{C}_z(z_b, t) \geq \mathcal{C}_z(z_s, t) .$$

Proof of Proposition 3.

In view of Proposition 4 and with Assumption H, we define partly the output function by letting :

$$h_z(z, t) = \mathcal{C}_z(z, t) \quad \forall (z, t) \in \{(z, t) \in \mathcal{D} : \mathcal{C}_z(z, t) \geq 0\} . \quad (25)$$

This incomplete definition is already sufficient for the implementation of the observer since it gives the function $z \mapsto \max\{0, h_z(z, t)\}$ or equivalently $\varphi \mapsto \max\{0, h(\varphi, t)\}$. But a complete definition is needed for the analysis. Namely we still need to define h_z outside the set :

$$\{(z, t) \in \mathcal{D} : \mathcal{C}_z(z, t) \geq 0\} .$$

For this, we take advantage of the remark made at the beginning of this Section, about the fact that nothing is imposed on pairs (φ, t) which are not in the set $\{(\varphi, t) : \mathcal{C}(\varphi, t) = 0\}$ and we rely on the convex function extension [15, Theorem 3.2]. However, because of the time dependence and our need to have bounds independent of t , we cannot use this result per se. Fortunately, we can follow step by step its proof.

To match the notations in [15, Proof of Theorem 3.2],

- we denote :

$$\varepsilon = \frac{1 - 2|e(t)|}{12} \geq \frac{\varrho}{6}$$

and, for k in $\{0, 1, 2, 3\}$,

$$\Omega^{k\varepsilon} = \left\{ (z, t) \in \mathcal{D} : \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 \geq \bar{\eta}(1 - |e(t)| - k\varepsilon) \right\} ,$$

knowing that we have :

$$\bar{\eta}(1 - |e(t)|) = 0 \quad , \quad \bar{\eta}(1 - |e(t)| - 3\varepsilon) = \bar{\eta} \left(\frac{3 - 2|e(t)|}{4} \right) \leq -\frac{\varrho}{12} < 0 .$$

With (25), the function h is already defined on Ω^0 and, from point 2 in Proposition 4, we have :

$$|z| \leq \frac{3}{2} \quad , \quad |e(t)| \leq \frac{1}{2} \quad \forall (z, t) \notin \Omega^0 . \quad (26)$$

Also, with point 3 of Proposition 4, for all (z, t) in $\Omega^\varepsilon \setminus \Omega^0$, we have :

$$d \left(z, \left\{ \mathfrak{z} \in \mathbb{R}^2 : |\mathfrak{z}|^2 - e(t)^\top \mathfrak{z} \leq \frac{3 - 2|e(t)|}{4} \right\} \right)$$

$$\begin{aligned}
&\geq \frac{1 - |e(t)| - \varepsilon - \frac{3-2|e(t)|}{4}}{2} = \frac{1 - |e(t)| - \frac{1-2|e(t)|}{12} - \frac{3-2|e(t)|}{4}}{2} = \frac{2 - 4|e(t)|}{24} = 2\varepsilon, \\
&\geq \frac{\varrho}{3}.
\end{aligned}$$

This implies, for all (z, t) in $\Omega^\varepsilon \setminus \Omega^0$,

$$d\left(\left\{\mathfrak{z} \in \mathbb{R}^2 : |\mathfrak{z} - z| \leq \varepsilon\right\}, \left\{\mathfrak{z} \in \mathbb{R}^2 : |\mathfrak{z}|^2 - e(t)^\top \mathfrak{z} \leq \frac{3-2|e(t)|}{4}\right\}\right) \geq \varepsilon \geq \frac{\varrho}{6}.$$

Hence, for any ζ with modulus smaller than δ in $[0, \varepsilon]$, $(z - \zeta, t)$ is in \mathcal{D} . Moreover with point 2 of Proposition 4, we have :

$$\mathcal{C}_z(z - \zeta, t) \geq \mathcal{C}_z(z, t) - 15\delta \geq \bar{\eta}(1 - |e(t)| - \varepsilon) - 15\delta.$$

Therefore, with Lemma 1, by letting :

$$\delta_{\max} = \min\left\{\varepsilon, \frac{\bar{\eta}(1 - |e(t)| - \varepsilon) - \bar{\eta}(1 - |e(t)| - 2\varepsilon)}{15}\right\} \geq \frac{\varrho}{6 * 3 * 15} > 0,$$

we get :

$$z - \zeta \in \Omega^{2\varepsilon} \quad \forall (z, \zeta) : (z, t) \in \Omega^\varepsilon \setminus \Omega^0, \quad |\zeta| \leq \delta_{\max}.$$

- With δ in $]0, \delta_{\max}[$, we let³ $p_\delta : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a C^2 function satisfying :

$$p_\delta(z) = 0 \quad \forall z : |z| \geq \delta, \quad \int_{\mathbb{R}^2} p_\delta(z) dz = 1. \quad (27)$$

It can be :

$$\begin{aligned}
p_\delta(z) &= \frac{945}{768} \frac{(|z|^2 - \delta^2)^4}{\delta^8} \quad \text{if } |z| \leq \delta, \\
&= 0 \quad \text{if } \delta < |z|.
\end{aligned}$$

- We let⁴ $\varpi : \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \rightarrow [0, 1]$ be a C^2 function satisfying :

$$\begin{aligned}
\varpi(z, t) &= 1 \quad \text{if } (z, t) \in \Omega^0, \\
&= 0 \quad \text{if } (z, t) \notin \Omega^\varepsilon.
\end{aligned} \quad (28)$$

It can be :

$$\begin{aligned}
\varpi(z) &= 1 \quad \text{if } z \in \Omega^0, \\
&= -(6x^2 + 3x + 1)(x - 1)^3 \quad \text{if } z \in \Omega^\varepsilon \setminus \Omega^0, \\
&= 0 \quad \text{if } z \notin \Omega^\varepsilon,
\end{aligned}$$

where :

$$x = \frac{\mathcal{C}_z(z, t)}{\bar{\eta}(1 - |e(t)| - \varepsilon)}.$$

With all this at hand, we define successively⁵ :

³ p_δ is denoted ϕ in [15, Proof of Theorem 3.2].

⁴ ϖ is denoted α in [15, Proof of Theorem 3.2].

⁵In [15, Proof of Theorem 3.2], \mathcal{C}_z is denoted f , $\bar{\mathcal{C}}_z$ is denoted g and h_z is denoted h .

1. The function $\bar{\mathcal{C}}_z$ as :

$$\begin{aligned}\bar{\mathcal{C}}_z(z, t) &= \mathcal{C}_z(z, t) && \text{if } (z, t) \in \Omega^{3\varepsilon}, \\ &= \bar{\eta} \left(\frac{3 - 2|e(t)|}{4} \right) && \text{if } (z, t) \notin \Omega^{3\varepsilon}.\end{aligned}$$

From 4 in Proposition 4 and the [15, Proof of Theorem 3.1], this gives a continuous function which is convex in z .

2. The function $\bar{\bar{\mathcal{C}}}_z$ as :

$$\bar{\bar{\mathcal{C}}}_z(z, t) = \int_{\mathbb{R}^2} \bar{\mathcal{C}}_z(\zeta, t) p_\delta(z - \zeta) d\zeta = \int_{\mathbb{R}^2} \bar{\mathcal{C}}_z(z - \zeta, t) p_\delta(\zeta) d\zeta. \quad (29)$$

From the [15, Proof of Theorem 3.2], this gives a C^2 function which is convex in z .

3. The function h_z as :

$$h_z(z, t) = \varpi(z) \mathcal{C}_z(z, t) + (1 - \varpi(z)) (\bar{\bar{\mathcal{C}}}_z(z, t) + c|z|^2), \quad (30)$$

where c is a strictly positive real number to be chosen. This gives a C^2 function. Because of (28), we have :m

$$\begin{aligned}h_z(z, t) &= \mathcal{C}_z(z, t) && \text{if } (z, t) \in \Omega^0, \\ &= \bar{\bar{\mathcal{C}}}_z(z, t) + c|z|^2 && \text{if } (z, t) \notin \Omega^\varepsilon.\end{aligned}$$

So (25) holds and, with (24), we get :

$$\begin{aligned}\frac{\partial^2 h_z}{\partial z \partial z}(z, t) &\geq 2\varphi I && \text{if } (z, t) \in \Omega^0, \\ &\geq cI && \text{if } (z, t) \notin \Omega^\varepsilon.\end{aligned}$$

On the other hand, because of (26), the function $z \mapsto \left(\mathcal{C}_z(z, t), \frac{\partial \mathcal{C}_z}{\partial z}(z, t), \frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z, t) \right)$ is Lipschitz uniformly on t outside the set Ω^0 . So, as in [15, Proof of Theorem 3.2], with (27) and the second expression in (29), it can be shown that δ can be chosen, independent of t , to make the difference $\bar{\bar{\mathcal{C}}}_z - \mathcal{C}_z$ as well as its first and second partial derivatives in z arbitrarily small on $\Omega^\varepsilon \setminus \Omega^0$. Then, by picking also c small enough, independently of t , we can obtain that the Hessian of $z \mapsto (1 - \varpi(z)) (\bar{\bar{\mathcal{C}}}_z(z, t) - \varpi(z) \mathcal{C}_z(z, t) + c|z|^2)$ is say smaller than φ for all t . This implies :

$$\frac{\partial^2 h_z}{\partial z \partial z}(z, t) \geq \min\{c, \varphi\} I \quad \forall (z, t).$$

□

We end this paragraph on the choice of the output function h or h_z by repeating that the only objective is to satisfy Assumptions A1, A2 and A3 where Assumption A3 is the only one linking the output function to the system. Above we have expressed this link by imposing that the level set $\{(z, t) : h_z(z, t) = 0\}$ be equal to the level set $\{(z, t) : z \neq 0, \mathcal{C}_z(z, t) = 0\}$. To define h_z outside this set, we have chosen to keep h_z as close as possible to \mathcal{C}_z . But this is only one possibility. There are others. We give one in Appendix C.

3.3 Total flux observer

As written above the total flux observer we propose is :

$$\dot{\hat{\varphi}} = v - Ri \quad \text{if} \quad \left[\left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi^2 |\hat{\varphi} - (L_0 - L_1) i(t)|^2 \right] \leq 0$$

and :

$$\begin{aligned} \dot{\hat{\varphi}} = v - Ri \\ + \mu \left[4 \left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right) (\hat{\varphi} - L_0 i(t)) - 2\Phi^2 (\hat{\varphi} - (L_0 - L_1) i(t)) \right] \times \\ \times \left[\left(|\hat{\varphi} - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi^2 |\hat{\varphi} - (L_0 - L_1) i(t)|^2 \right] \end{aligned}$$

if not. Let $\hat{\varphi}(\hat{\varphi}_0, t)$ denote its solution at time t , issued from $\hat{\varphi}_0$ at time 0. Also, we say the function $t \mapsto \varphi_s(\varphi_0, t)$ is a constrained solution of the motor if it is a solution at time t of

$$\dot{\varphi} = v(t) - Ri(t) \quad (31)$$

going through φ_0 at time 0 and such that, for each t , there exists $\theta(t)$, called electrical angular position, satisfying :

$$\varphi_s(\varphi_0, t) = L_0 i(t) + \Phi \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} + L_1 \begin{pmatrix} \cos(2\theta(t)) & \sin(2\theta(t)) \\ \sin(2\theta(t)) & -\cos(2\theta(t)) \end{pmatrix} i(t) . \quad (32)$$

Proposition 5. *Assume :*

1. $v - Ri$ as well as its integral is bounded on $\mathbb{R}_{\geq 0}$;
2. there exists a strictly positive real number \wp in $\left] 0, \frac{1}{2} \right]$ such that we have :

$$2 \frac{L_1}{\Phi} |i(t)| \leq \frac{1}{2} - \wp \quad \forall t \geq 0 . \quad (33)$$

3. The time derivative of θ in (32) is not integrable.

Then, for any μ strictly positive and any $(\varphi_0, \hat{\varphi}_0)$, source of a constrained solution, we have :

$$\left| \varphi_s(\varphi_0, t) - \hat{\varphi}(\hat{\varphi}_0, t) \right| \leq |\varphi_0 - \hat{\varphi}_0| \quad \forall t \geq 0$$

and

$$\lim_{t \rightarrow \infty} \left| \varphi_s(\varphi_0, t) - \hat{\varphi}(\hat{\varphi}_0, t) \right| = 0 .$$

Proof. With the output function h given by Proposition 3, φ playing the role of x and $v - Ri$, the role of u , a constrained solution $\varphi_s(\varphi_0, t)$ of (31) and (32) is a constrained solution of (6) and (7). Moreover the observer (14) is nothing but the counter-part of (10). Hence the result follows directly from Proposition 1 provided assumptions A1, A2 and A3 are satisfied. Proposition 3 guarantees this is the case for assumptions A1 and A2. So it remains only to check that assumption A3 is satisfied.

It follows from Proposition 2 and (33) that $\varphi_s(\varphi_0, t)$ satisfies :

$$\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t) \neq 0 \quad (34)$$

and

$$\left(|\varphi_s(\varphi_0, t) - L_0 i(t)|^2 - |L_1 i(t)|^2 \right)^2 - \Phi(t)^2 |\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)|^2 = 0$$

or equivalently

$$\begin{aligned} \left(|\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)|^2 - 2L_1 i(t)^\top [\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)] \right)^2 \\ = \Phi(t)^2 |\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)|^2 . \end{aligned} \quad (35)$$

On another hand, (32) implies :

$$\varphi_s(\varphi_0, t) - L_0 i(t) + L_1 i(t) = \left(\Phi + 2L_1(\cos(\theta(t)) \sin(\theta(t))) i(t) \right) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} .$$

So the argument of $\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)$ is the electrical angular position $\theta(t)$. Let ρ be its modulus. Let also λ and σ be the modulus and argument of $2L_1 i(t)$, i.e.

$$\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t) = \rho(t) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} , \quad 2L_1 i(t) = \lambda(t) \begin{pmatrix} \cos(\sigma(t)) \\ \sin(\sigma(t)) \end{pmatrix} .$$

Note that (33) implies

$$\lambda(t) < \frac{\Phi}{2} . \quad (36)$$

Equations (35) and (34) become :

$$\left(\rho(t)^2 - \rho(t)\lambda(t) \cos(\sigma(t) - \theta(t)) \right)^2 = \Phi^2 \rho(t)^2 , \quad \rho(t) \neq 0 .$$

With (36), this implies :

$$\rho(t) = \lambda(t) \cos(\sigma(t) - \theta(t)) + \Phi .$$

On another hand, we have :

$$\begin{aligned} \frac{\partial h}{\partial \varphi}(\varphi_s(\varphi_0, t), t) \\ = 4 \left(|\varphi_s(\varphi_0, t) - L_0 i(t)|^2 - |L_1 i(t)|^2 \right) (\varphi_s(\varphi_0, t) - L_0 i(t))^\top - 2\Phi^2 (\varphi_s(\varphi_0, t) - (L_0 - L_1) i(t)) , \\ = 4 \left(\rho(t)^2 - \rho(t)\lambda(t) \cos(\sigma(t) - \theta(t)) \right) \left(\rho(t) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} - \frac{\lambda(t)}{2} \begin{pmatrix} \cos(\sigma(t)) \\ \sin(\sigma(t)) \end{pmatrix} \right) \\ \quad - 2\Phi^2 \rho(t) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} , \\ = 4\Phi \rho(t) \left(\rho(t) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} - \frac{\lambda(t)}{2} \begin{pmatrix} \cos(\sigma(t)) \\ \sin(\sigma(t)) \end{pmatrix} \right) - 2\Phi^2 \rho(t) \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} , \\ = 2\Phi \rho(t) \begin{pmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{pmatrix} \begin{pmatrix} \lambda(t) \cos(\sigma(t) - \theta(t)) + \Phi \\ -\lambda(t) \sin(\sigma(t) - \theta(t)) \end{pmatrix} . \end{aligned}$$

Then let $w = (\cos(\vartheta) \ \sin(\vartheta))^\top$ be an arbitrary unit vector. We obtain :

$$\begin{aligned} \frac{\partial h}{\partial \varphi}(\varphi_s(\varphi_0, t), t) w \\ = 2\Phi \rho(t) \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \end{pmatrix} \begin{pmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{pmatrix} \begin{pmatrix} \lambda(t) \cos(\sigma(t) - \theta(t)) + \Phi \\ -\lambda(t) \sin(\sigma(t) - \theta(t)) \end{pmatrix} , \\ = 2\Phi \rho(t) (\cos(\vartheta - \theta(t)) \ \sin(\vartheta - \theta(t))) \begin{pmatrix} \lambda(t) \cos(\sigma(t) - \theta(t)) + \Phi \\ -\lambda(t) \sin(\sigma(t) - \theta(t)) \end{pmatrix} . \end{aligned}$$

Since $\dot{\theta}$ is not integrable, there exists a sequence t_n going to infinity as n grows indefinitely such that :

$$\theta(t_n) = \vartheta .$$

This yields :

$$\frac{\partial h}{\partial \varphi}(\varphi_s(\varphi_0, t_n), t_n)w = \lambda(t) \cos(\sigma(t_n) - \theta(t_n)) + \Phi > \frac{\Phi}{2} .$$

So assumption A3 holds. □

3.4 Rotor position estimation

To obtain a sensorless control for salient-pole synchronous machines, we can proceed exactly as in [8, 6, 9]. It is sufficient for this to derive an estimate $\hat{\theta}$ of the electrical angular position θ from the estimate $\hat{\varphi}$ of the total flux.

Depending on the online implementation constraints and controllers' resources, one of the following method can be used.

3.5 Method 1

We know with Proposition 2 that, if the total flux satisfies :

$$\mathcal{C}(\varphi, t) = 0 , \tag{37}$$

then we have (18). Since, in an ideal context - perfect model, no noise, . . . - , Proposition 5 insures the total flux estimate $\hat{\varphi}$ satisfies this equation asymptotically, a first method is :

$$\begin{pmatrix} \cos(\hat{\theta}) \\ \sin(\hat{\theta}) \end{pmatrix} = \frac{\hat{\varphi} - (L_0 - L_1)i(t)}{|\hat{\varphi} - (L_0 - L_1)i(t)|} .$$

3.6 Method 2

Total flux and electrical angular position are related by (2). With denoting :

$$\varphi = (\varphi_\alpha \ \varphi_\beta)^\top , \quad i = (i_\alpha \ i_\beta)^\top$$

and making use of some trivial trigonometric relations, (2) leads to the following system of polynomial equations in $c = \cos(\theta)$ and $s = \sin(\theta)$:

$$\begin{aligned} 2L_1 i_\beta c s + 2L_1 i_\alpha c^2 + \Phi c - \varphi_\alpha + (L_0 - L_1)i_\alpha &= 0 , \\ 2L_1 i_\alpha c s - 2L_1 i_\beta c^2 + \Phi s - \varphi_\beta + (L_0 + L_1)i_\beta &= 0 , \\ c^2 + s^2 &= 1 . \end{aligned} \tag{38}$$

With ignoring the last trigonometric identity, the existence of at least one real solution to the first two equations of this system is guaranteed. To see this, we eliminate s from the above two equations. This yields :

$$2L_1 i_\beta c (-2L_1 i_\beta c^2 - \varphi_\beta + (L_0 + L_1)i_\beta) = (2L_1 i_\alpha c + \Phi) (2L_1 i_\alpha c^2 + \Phi c - \varphi_\alpha + (L_0 - L_1)i_\alpha) .$$

So we end up with a polynomial of degree 3 in c :

$$\begin{aligned} 4L_1^2 |i|^2 c^3 + 4L_1 i_\alpha \Phi c^2 \\ + (\Phi^2 - 2L_1^2 |i|^2 - 2L_1 L_0 (i_\alpha^2 - i_\beta^2) + 2L_1 (i_\beta \varphi_\beta - i_\alpha \varphi_\alpha)) c - \Phi (\varphi_\alpha - (L_0 - L_1)i_\alpha) . \end{aligned}$$

It has at least one real solution and all its root can be obtained via Cardan's Method. Afterwards, the corresponding s is determined from (38) as :

$$s = -\frac{(2L_1 i_\beta c)(2L_1 i_\alpha c^2 + \Phi c - \varphi_\alpha + (L_0 - L_1)i_\alpha) + (2L_1 i_\alpha c + \Phi)(-2L_1 i_\beta c^2 - \varphi_\beta + (L_0 + L_1)i_\beta)}{(2L_1 i_\beta c)^2 + (2L_1 i_\alpha c + \Phi)^2}.$$

Not knowing the actual flux φ , we replace it by its estimate $\hat{\varphi}$ given by the observer. Proceeding as above after this substitution gives us up to 3, but at least one, real estimates of pairs (\hat{c}, \hat{s}) . Among the real ones, say (\hat{c}_1, \hat{s}_1) is the one the norm of which is closest to 1. Then an estimate of the electrical angular position is obtained from :

$$\begin{pmatrix} \cos(\hat{\theta}) \\ \sin(\hat{\theta}) \end{pmatrix} = \frac{\begin{pmatrix} \hat{c}_1 \\ \hat{s}_1 \end{pmatrix}}{\sqrt{\hat{c}_1^2 + \hat{s}_1^2}}.$$

3.7 Method 3

Acknowledging that the estimate $\hat{\varphi}$ possibly does not satisfy (37) nor (2), we consider the following least square problem :

$$(\hat{c}, \hat{s}) = \arg \min_{(c,s):c^2+s^2=1} \left| \hat{\varphi} - L_0 i - \Phi \begin{pmatrix} c \\ s \end{pmatrix} - L_1 \begin{pmatrix} 2c^2 - 1 & 2sc \\ 2sc & 1 - 2c^2 \end{pmatrix} i \right|^2.$$

Then, determining a solution (\hat{c}, \hat{s}) has to to be performed numerically online.

4 Image of the unit circle by a polynomial function

If, instead of (2), the flux satisfies the more general expression (5), an observer can still be given by (10) as long as we know the flux time derivative $\dot{\varphi}$ and provided we can express (10) as an algebraic constraint $\mathcal{C}(\varphi, t) = 0$. Synchronous machines somehow fall under this scope, first because they are fully described by (1), and second because there exists such a constraint that may be analytically computed even when the total flux is expressed as (5). In the following, we give a systematic method to derive the corresponding constraint function \mathcal{C} .

As in section 3.1, we use the complex representations and adopt the following notations :

- $\varphi = \varphi_\alpha + \mathbf{j}\varphi_\beta \in \mathbb{C}$ is the complex representation of the total flux vector $\varphi = (\varphi_\alpha \ \varphi_\beta)^\top$,
- $\xi = \cos(\theta) + \mathbf{j}\sin(\theta)$,
- $\gamma_k = \varphi_k [\cos(\psi_k) + \mathbf{j}\sin(\psi_k)]$.

As in section 3.1, the fact that φ satisfies (5) is equivalent to the fact that the polynomial

$$P(\xi) = (\gamma_0 - \varphi) + \gamma_1 \xi + \gamma_2 \xi^2 + \dots + \gamma_N \xi^N.$$

has a root on the unit circle.

With \mathbf{A} and \mathbf{B} denoting the following Toeplitz matrices :

$$\mathbf{A} = \begin{pmatrix} \gamma_0 - \varphi & \gamma_1 & \dots & \dots & \gamma_{N-1} \\ 0 & \gamma_0 - \varphi & \gamma_1 & & \gamma_{N-2} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \gamma_1 \\ 0 & \dots & \dots & 0 & \gamma_0 - \varphi \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \gamma_N & 0 & \dots & \dots & 0 \\ \gamma_{N-1} & \gamma_N & \ddots & & \vdots \\ \vdots & \gamma_{N-1} & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \gamma_1 & \dots & \dots & \gamma_{N-1} & \gamma_N \end{pmatrix}. \quad (39)$$

we have the following result.

Proposition 6. *If φ is given by (5), then, for each t , we have :*

$$\mathcal{C}(\varphi, t) = 0 ,$$

where $\mathcal{C}(\varphi, t)$ is a polynomial of degree $2N$ in the components $(\varphi_\alpha, \varphi_\beta)$ of φ , with real coefficients and leading term $(\varphi_\alpha^2 + \varphi_\beta^2)^N$, given by :

$$\mathcal{C}(\varphi, t) = \det \left(\mathbf{A}\overline{\mathbf{A}}^\top - \overline{\mathbf{B}}^\top \mathbf{B} \right) . \quad (40)$$

Proof. The proof goes exactly the same way as the proof of Proposition 2. Namely if ξ is a root of P with modulus 1, then it is also a root of :

$$Q(\xi) = \overline{\gamma}_N + \overline{\gamma}_{N-1} \xi + \dots + \overline{\gamma}_1 \xi^{N-1} + (\overline{\gamma}_0 - \overline{\varphi}) \xi^N .$$

So P and Q have a common root which is equivalent to :

$$\det(M) = 0$$

where :

$$M = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \overline{\mathbf{B}}^\top & \overline{\mathbf{A}}^\top \end{pmatrix} .$$

There A and B are Toeplitz matrices which, because of

$$\sum_{\ell=i}^j \overline{\gamma}_{N+i-\ell} \gamma_{j-\ell} = \sum_{k=i}^j \gamma_{k-i} \overline{\gamma}_{N+k-j}$$

satisfy :

$$\mathbf{A}\overline{\mathbf{B}}^\top = \overline{\mathbf{B}}^\top \mathbf{A}$$

and therefore :

$$\begin{pmatrix} I & 0 \\ -\overline{\mathbf{B}}^\top & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \overline{\mathbf{B}}^\top & \overline{\mathbf{A}}^\top \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{A}\overline{\mathbf{A}}^\top - \overline{\mathbf{B}}^\top \mathbf{B} \end{pmatrix} ,$$

$$\begin{pmatrix} \overline{\mathbf{A}}^\top & -\mathbf{B} \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \overline{\mathbf{B}}^\top & \overline{\mathbf{A}}^\top \end{pmatrix} = \begin{pmatrix} \overline{\mathbf{A}}^\top \mathbf{A} - \mathbf{B}\overline{\mathbf{B}}^\top & 0 \\ \overline{\mathbf{B}}^\top & \overline{\mathbf{A}}^\top \end{pmatrix} .$$

So we have algebraically :

$$\det(M) = \det \left(\mathbf{A}\overline{\mathbf{A}}^\top - \overline{\mathbf{B}}^\top \mathbf{B} \right) = \det \left(\overline{\mathbf{A}}^\top \mathbf{A} - \mathbf{B}\overline{\mathbf{B}}^\top \right) = \overline{\det(M)} .$$

Hence $\det(M)$ is real. It is an homogeneous polynomial of degree $2N$ in $(\varphi - \gamma_0, \gamma_1, \dots, \gamma_N)$ and a polynomial, with real coefficients, of degree $2N$ in the components of φ . Also, given the multilinearity of the determinant function, the higher degree term in $\varphi - \gamma_0$, respectively in φ , is $|\varphi - \gamma_0|^{2N}$, $|\varphi|^{2N}$ respectively. \square

To study the properties of the constraint function (40), we assume that γ_1 is not zero. In the case where γ_2 to γ_N are zero, we obtain :

$$\mathcal{C}(\varphi, t)$$

$$\begin{aligned}
&= \det \left(\begin{array}{cccc|cc|cccc}
-\varphi + \gamma_0 & \gamma_1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & \ddots & \ddots & \ddots & & & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & & & & & \vdots \\
\vdots & & 0 & -\varphi + \gamma_0 & \gamma_1 & 0 & & & & \vdots \\
\hline
\vdots & & & 0 & -\varphi + \gamma_0 & \gamma_1 & 0 & & & \vdots \\
\vdots & & & 0 & \bar{\gamma}_1 & -\bar{\varphi} + \bar{\gamma}_0 & 0 & & & \vdots \\
\hline
\vdots & & & & 0 & \bar{\gamma}_1 & -\bar{\varphi} + \bar{\gamma}_0 & \ddots & & \vdots \\
\vdots & & & & & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & & & \ddots & \ddots & \ddots & 0 \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \bar{\gamma}_1 & -\bar{\varphi} + \bar{\gamma}_0
\end{array} \right) \\
&= |\varphi - \gamma_0|^{2(N-1)} (|\varphi - \gamma_0|^2 - |\gamma_1|^2).
\end{aligned}$$

Hence, in this case, for $N > 1$, \mathcal{C} is zero if and only if either :

$$\varphi = \varphi_0 \begin{pmatrix} \cos(\psi_0) \\ \sin(\psi_0) \end{pmatrix}$$

or φ is in the circle with center $\varphi_0 \begin{pmatrix} \cos(\psi_0) \\ \sin(\psi_0) \end{pmatrix}$ and radius $|\gamma_1|$.

When γ_2 to γ_N are not zero but small in modulus, we let :

$$\rho [\cos(\varpi) + j \sin(\varpi)] = \varphi - \gamma_0 \quad , \quad \varepsilon = \sqrt{\sum_{k=2}^N |\gamma_k|^2} \quad , \quad \underline{\gamma}_k = \frac{\gamma_k}{\varepsilon} .$$

In the general case, the equation $\mathcal{C} = 0$ can be rewritten :

$$\rho^{2N} - \rho^{2(N-1)} |\gamma_1|^2 + \varepsilon \sum_{k=0}^{2N-1} \rho^k \varepsilon^{2N-1-k} c_k \left(\cos(\varpi), \sin(\varpi), \underline{\gamma}_2, \dots, \underline{\gamma}_N \right) = 0 ,$$

where the c_k are polynomials in their arguments which are bounded. It follows from Rouché Theorem that, for any strictly positive real number $r < |\gamma_1|$, there exists a strictly positive real number ε_* such that, for all ε in $[0, \varepsilon^*]$, all ϖ in $[0, 2\pi[$ and all $N - 1$ -uple of complex numbers $\underline{\gamma}_k$ satisfying :

$$\sum_{k=2}^N |\underline{\gamma}_k|^2 = 1, \tag{41}$$

the above equation in ρ has $2N$ complex zeros, one in the ball with center $|\gamma_1|$ and radius r , and therefore real and strictly positive, one in the ball with center $-|\gamma_1|$ and radius r and therefore real and strictly negative, the other $2(N - 1)$ in the the ball centered at the origin and radius r . Since we are interested only with ρ non negative, we focus on the root in the ball with center $|\gamma_1|$ and radius r . It can be expressed as :

$$\rho = R \left(\cos(\varpi), \sin(\varpi), \underline{\gamma}_2, \dots, \underline{\gamma}_N \right) ,$$

where :

$$\left| R \left(\cos(\varpi), \sin(\varpi), \underline{\gamma}_2, \dots, \underline{\gamma}_N \right) - |\gamma_1| \right| \leq r .$$

By letting $(\cos(\varpi), \sin(\varpi))$ go around the unit circle, we obtain the counterpart of the Limaçon of Pascal.

Appendices

A Proof of Lemma 1

We study the properties of the function $\eta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ defined as :

$$\eta(\bar{h}) = \bar{h}^2 - \frac{1}{4} \left[e - \sqrt{e^2 + 4\bar{h}} \right]^2$$

where e is a given real number in $\left[0, \frac{1}{2} - \wp\right]$. Its graph is shown on Figure 1

Before this, it is useful to note that, actually from its construction given in Section B.3, the function $\bar{\eta}$ is the composition of the function $\bar{h} \mapsto \frac{e}{2} - \sqrt{\frac{e^2}{4} + \bar{h}}$ and the function $z \mapsto (z^2 - ez)^2 - z^2$, i.e.

$$\begin{aligned} \bar{\eta}(\bar{h}) &= (z^2 - ez)^2 - z^2 \Big|_{z=\frac{e}{2}-\sqrt{\frac{e^2}{4}+\bar{h}}} , \\ &= z^2[(z-e)^2 - 1] \Big|_{z=\frac{e}{2}-\sqrt{\frac{e^2}{4}+\bar{h}}} . \end{aligned}$$

Since the function $z \mapsto (z-e)^2 - 1$ is strictly negative for all z in the open interval $]e-1, 0[$, and :

$$e-1 = \frac{e}{2} - \sqrt{\frac{e^2}{4} + (1-e)} \quad , \quad 0 = \frac{e}{2} - \sqrt{\frac{e^2}{4}} \quad ,$$

a direct consequence is :

$$\bar{\eta}(\bar{h}) < 0 \quad \forall \bar{h} \in]0, 1-e[. \quad (42)$$

The first derivative of $\bar{\eta}$ is :

$$\begin{aligned} \frac{d\bar{\eta}}{d\bar{h}}(\bar{h}) &= 2\bar{h} + \frac{1}{2} \left[e - \sqrt{e^2 + 4\bar{h}} \right] \frac{4}{2\sqrt{e^2 + 4\bar{h}}} , \\ &= 2\bar{h} - 1 + \frac{e}{\sqrt{e^2 + 4\bar{h}}} . \end{aligned}$$

It is zero on $\mathbb{R}_{\geq 0}$ if and only if :

$$\bar{h} = 0 \quad \text{or} \quad = \frac{4 - e^2 - e\sqrt{e^2 + 8}}{8} .$$

According to (45) the former case is given by the local maximizer $\bar{z} = 0$ of the function \mathcal{C}_z , while the latter is given by the saddle point \bar{z} . This implies $\bar{\eta}$ is non increasing on the closed interval $\left[0, \frac{4 - e^2 - e\sqrt{e^2 + 8}}{8}\right]$ and strictly increasing on the open interval $\left[\frac{4 - e^2 - e\sqrt{e^2 + 8}}{8}, +\infty\right[$.

The second derivative is :

$$\frac{d^2\bar{\eta}}{d\bar{h}^2}(\bar{h}) = 2 \left(1 - \frac{e}{(e^2 + 4\bar{h})^{3/2}} \right)$$

It is negative on $\left[0, \frac{e^{2/3} - e^2}{4}\right]$ and strictly positive on $\left[\frac{e^{2/3} - e^2}{4}, +\infty\right]$. This implies $\bar{\eta}$ is strictly convex on $\left[\frac{e^{2/3} - e^2}{4}, +\infty\right]$, its first derivative is strictly increasing on this interval and

$$\bar{\eta}(\bar{h}_a) \geq \bar{\eta}(\bar{h}_b) + \frac{d\bar{\eta}}{d\bar{h}}(\bar{h}_b)(\bar{h}_a - \bar{h}_b) = \left(2\bar{h}_b - 1 + \frac{e}{\sqrt{e^2 + 4\bar{h}_b}}\right)(\bar{h}_a - \bar{h}_b) \quad \forall \bar{h}_a, \bar{h}_b \geq \frac{e^{2/3} - e^2}{4}. \quad (43)$$

To have the 3 properties (42), monotonicity of $\frac{d\eta}{d\bar{h}}$ and (43) on the same interval, it is sufficient to pick a bound between $\frac{4 - e^2 - e\sqrt{e^2 + 8}}{8}$ and $1 - e$. We choose $\frac{3 - 2e}{4}$. We have established that $\bar{\eta}$ and its derivative are strictly increasing on $\left[\frac{3 - 2e}{4}, +\infty\right]$ and satisfies :

$$\bar{\eta}(\bar{h}_a) \geq \bar{\eta}(\bar{h}_b) + \left(2\bar{h}_b - 1 + \frac{e}{\sqrt{e^2 + 4\bar{h}_b}}\right)(\bar{h}_a - \bar{h}_b) \geq \frac{1}{3}(\bar{h}_a - \bar{h}_b) \quad \forall \bar{h}_a, \bar{h}_b \geq \frac{3 - 2e}{4}.$$

So we have in particular :

$$0 = \bar{\eta}(1 - e) \geq \bar{\eta}\left(\frac{3 - 2e}{4}\right) + \frac{1 - e - \frac{3 - 2e}{4}}{3} = \bar{\eta}\left(\frac{3 - 2e}{4}\right) + \frac{1 - 2e}{12} \geq \bar{\eta}\left(\frac{3 - 2e}{4}\right) + \frac{\wp}{12}.$$

B Properties of the constraint function (15)

B.1 About point 1 in Proposition 4

The equation :

$$0 = \mathfrak{C}_z(z, t) = \left(|z|^2 - e(t)^\top z\right)^2 - |z|^2$$

is the Cartesian form of a Limaçon of Pascal. A lot is known about this curve. See [7, pp. 44-51] for example. Its shape depends on the position of $|e|$ with respect to $\frac{1}{2}$ and 1. In particular, (see Figure 2)

- for $1 < |e(t)|$ it is not simple (it has two loops),
- for $\frac{1}{2} < |e(t)| \leq 1$ it encircles a non (Euclidean-) convex set,
- for $|e(t)| \leq \frac{1}{2}$ it encircles a (Euclidean-) convex set.

So, knowing from point 2 of Remark 1 that the sub-level set $\{\varphi : h(\varphi, t) \leq 0\}$, where the actual total flux evolves, should be convex and since the change of coordinates in (21) is affine, we impose⁶ the existence of a strictly positive real number \wp such that we have :

$$\left(2\frac{L_1}{\Phi}|i(t)| = \right) |e(t)| \leq \frac{1}{2} - \wp \quad \forall t > 0.$$

This says the saliency or the current should be small.

⁶This constraint is imposed to obtain an Euclidean convexity. It can be relaxed if geodesic convexity for some appropriately chosen metric is considered.

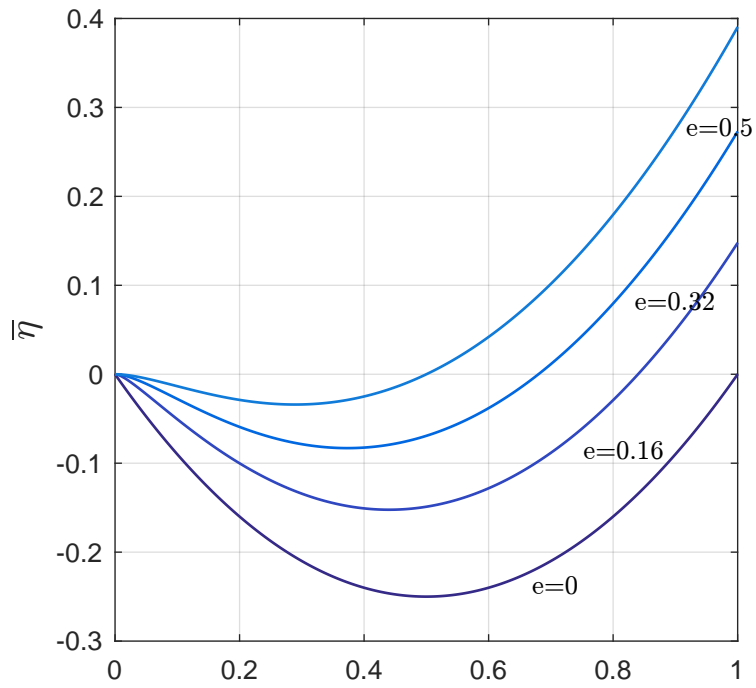


Figure 1: Function $\bar{\eta}$ for various values of $|e|$

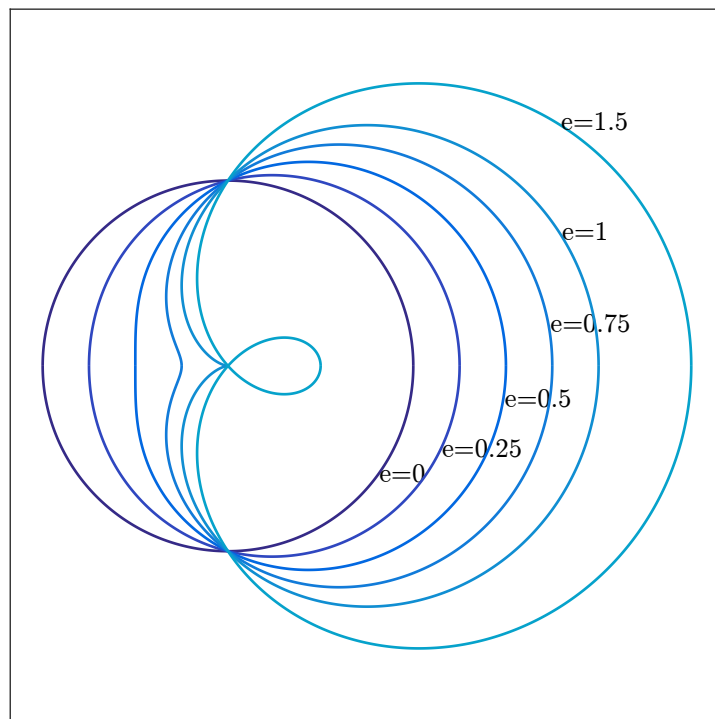


Figure 2: Limaçon of Pascal for various values of $|e|$

B.2 About point 2 in Proposition 4

Let (z, t) satisfy :

$$0 \geq \mathcal{C}_z(z, t) = \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 .$$

We obtain :

$$|z|^2 - |e(t)| |z| \leq |z|^2 - e(t)^\top z \leq \left| |z|^2 - e(t)^\top z \right| \leq |z| .$$

With Assumption H, this implies :

$$|z| \leq 1 + |e(t)| \leq \frac{3}{2} .$$

Let (z_a, z_b, t) be such that :

$$\mathcal{C}_z(z_a, t) \leq 0 \quad , \quad \mathcal{C}_z(z_b, t) \leq 0 .$$

Since, under Assumption H, for each t , the set $\{z : \mathcal{C}_z(z, t) \leq 0\}$ is convex, we have :

$$\mathcal{C}_z(z_s, t) \leq 0 ,$$

where :

$$z_s = z_b + s(z_a - z_b) \quad s \in [0, 1] .$$

This yields :

$$\begin{aligned} \left| \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t) \right| &= 2 \left| \left(|z_s|^2 - e(t)^\top z_s \right) (2z_s - e(t))^\top - z_s^\top \right| , \\ &\leq 2 \left(\left| |z_s|^2 - e(t)^\top z_s \right| (2|z_s| + |e(t)|) + |z_s| \right) , \\ &\leq 2 (|z_s| (2|z_s| + |e(t)|) + |z_s|) , \\ &\leq 2(1 + |e(t)|)(3 + 4|e(t)|) \leq 15 . \end{aligned}$$

So we get :

$$\begin{aligned} |\mathcal{C}_z(z_a, t) - \mathcal{C}_z(z_b, t)| &= \left| \int_0^1 \frac{\partial \mathcal{C}_z}{\partial z}(z_b + s[z_a - z_b]) ds [z_a - z_b] \right| , \\ &\leq \sup_{s \in [0, 1]} \left| \frac{\partial \mathcal{C}_z}{\partial z}(z_b + s[z_a - z_b]) \right| |z_a - z_b| , \\ &\leq 15 |z_a - z_b| . \end{aligned}$$

B.3 About point 3 in Proposition 4

To obtain information on the level sets of the constraint function \mathcal{C}_z , we compute :

$$\frac{\partial \mathcal{C}}{\partial \varphi}(\varphi, t) = \frac{1}{\Phi} \frac{\partial \mathcal{C}_z}{\partial z}(z, t) = \frac{2}{\Phi} \left(\left(|z|^2 - e(t)^\top z \right) (2z - e(t))^\top - z(t)^\top \right) ,$$

$$\frac{\partial^2 \mathcal{C}}{\partial \varphi \partial \varphi}(\varphi, t) = \frac{1}{\Phi^2} \frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z, t) = \frac{2}{\Phi^2} \left((2z - e(t)) (2z - e(t))^\top + 2 \left(|z|^2 - e(t)^\top z \right) - 1 \right) .$$

Hence the eigen vectors of the Hessian are $(2z - e(t))$ and its orthogonal. So its eigen values are :

$$\frac{2}{\Phi^2} \left(|2z - e(t)|^2 + 2 \left(|z|^2 - e(t)^\top z \right) - 1 \right) \quad , \quad \frac{2}{\Phi^2} \left(2 \left(|z|^2 - e(t)^\top z \right) - 1 \right) . \quad (44)$$

The gradient is zero if and only if z is colinear with $e(t)$, i.e. we have :

$$z = \zeta e(t) ,$$

with ζ satisfying :

$$0 = \zeta [|e(t)|^2 (\zeta - 1) (2\zeta - 1) - 1] .$$

Therefore the gradient is zero at :

$$\begin{aligned} \bar{z} &= z_1 = \zeta_1 e(t) = 0 , \\ \underline{z} &= z_2 = \zeta_2 e(t) = \frac{3|e(t)| + \sqrt{8 + |e(t)|^2}}{4} \frac{e}{|e|} , \\ \bar{z} &= z_3 = \zeta_3 e(t) = \frac{3|e(t)| - \sqrt{8 + |e(t)|^2}}{4} \frac{e}{|e|} . \end{aligned} \tag{45}$$

We obtain :

$$\mathcal{C}_z(\bar{z}, t) = 0 \quad , \quad \mathcal{C}_z(\underline{z}, t) < 0 \quad , \quad \mathcal{C}_z(\bar{z}, t) < 0 .$$

The Hessian at these stationary points is :

$$\frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z_i, t) = 2 \left((2\zeta_i - 1)^2 e(t) e(t)^\top + 2 (\zeta_i^2 - \zeta_i) |e(t)|^2 - 1 \right) .$$

Its eigen values are :

$$-2(1 - |e(t)|^2) \quad , \quad -2 ,$$

for ζ_1 , and :

$$\begin{aligned} 2 \left((2\zeta_i - 1)^2 + 2 (\zeta_i^2 - \zeta_i) \right) |e(t)|^2 - 2 &= 2 (2 + (3\zeta_i - 2) |e(t)|^2) , \\ 4 (\zeta_i^2 - \zeta_i) |e(t)|^2 - 2 &= 2 (\zeta_i - 1) |e(t)|^2 , \end{aligned}$$

for ζ_2 and ζ_3 . We conclude the constraint function \mathcal{C}_z has 3 stationary points, all in the sub-level set $\{z : \mathcal{C}_z(z, t) \leq 0\}$, a global strict minimum at \underline{z} , a local strict maximum at $\bar{z} = 0$, and a saddle point at \bar{z} . It follows that, for η in $]\mathcal{C}_z(\bar{z}, t), \mathcal{C}_z(\bar{z}, t)[$, the super-level set $\{z \in \mathbb{R}^2 : (|z|^2 - e(t)^\top z)^2 - |z|^2 \geq \eta\}$ have two connected components. See figure 3.

With Proposition 2, we know that the solution $z = 0$ to $|z|^2 - e(t)^\top z)^2 - |z|^2 = 0$ should be disregarded. In other words, we can eliminate a neighborhood of the connected component containing $z = 0$. To do so we restrict the domain of the constraint function \mathcal{C}_z and correspondingly of \mathcal{C} by removing the set :

$$\mathcal{S}_{ex} = \left\{ (z, t) : |z|^2 - e(t)^\top z \leq \underline{\mathfrak{h}} \right\} ,$$

where the bound $\underline{\mathfrak{h}}$ is chosen as :

$$\underline{\mathfrak{h}} = \frac{3 - 2|e(t)|}{4} .$$

With t fixed, this is the disk with center $\frac{e(t)}{2}$ and radius $\sqrt{\underline{\mathfrak{h}} + \frac{|e(t)|^2}{4}}$, in grey in Figure 3. It contains the local maximizer $\bar{z} = 0$ and the saddle point \bar{z} .

From now on, the domain of definition of \mathcal{C}_z is considered as being :

$$\mathcal{D} = \left\{ (z, t) \in \mathbb{R}^2 \times \mathbb{R}_{\geq 0} : |z|^2 - e(t)^\top z > \underline{\mathfrak{h}} \right\} .$$

Then, by super-, sub-, level set of \mathcal{C}_z , we mean :

$$\{(z, t) \in \mathcal{D} : \mathcal{C}_z(z, t) > (\leq) (=) \eta\} .$$

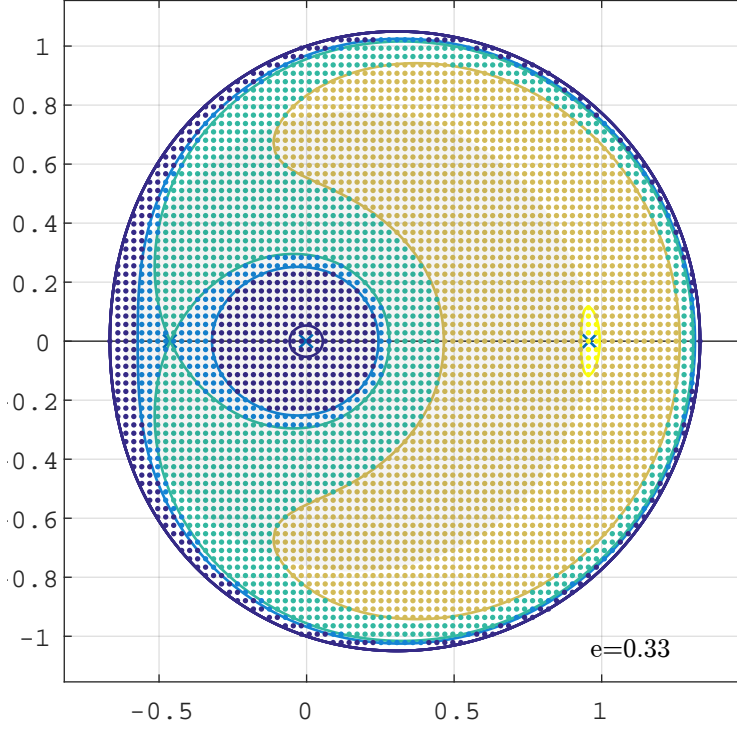


Figure 3: Level sets of \mathcal{C}_z with the excluded disk in grey

It follows from the above analysis, and (44) in particular, that, for all (z, t) in \mathcal{D} , the corresponding Hessian, in z , of \mathcal{C}_z satisfies :

$$\frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z, t) \geq 2 \left(2 \left(|z|^2 - e(t)^\top z \right) - 1 \right) \geq 2(2\underline{h} - 1) I = (1 - 2|e(t)|) I \geq 2\wp I . \quad (46)$$

To continue our analysis of the super-level sets of \mathcal{C}_z , we address the the following maximization problem :

$$\bar{\eta}(\bar{h}) = \sup_{z: \underline{h} \leq |z|^2 - e(t)^\top z \leq \bar{h}} \left\{ \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 \right\} , \quad (47)$$

with \bar{h} arbitrary in $[\underline{h}, 1 - |e(t)|]$. Our motivation is that, if (z, t) satisfies :

$$\left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 > \bar{\eta}(\bar{h}) ,$$

then we have :

- either $|z|^2 - e(t)^\top z < \underline{h}$. In this case (z, t) is not in \mathcal{D} .
- or $|z|^2 - e(t)^\top z > \bar{h}$. In this case we have (46) and, by comparing the radius of the disks, we see that the distance $d\left(z, \left\{ \zeta \in \mathbb{R}^2 : |\zeta|^2 - e(t)^\top \zeta \leq \underline{h} \right\}\right)$ between z and the excluded

disk $\left\{ \zeta \in \mathbb{R}^2 : |\zeta|^2 - e(t)^\top \zeta \leq \underline{\mathfrak{h}} \right\}$ is larger or equal to :

$$\begin{aligned} \sqrt{\bar{\mathfrak{h}} + \frac{|e(t)|^2}{4}} - \sqrt{\underline{\mathfrak{h}} + \frac{|e(t)|^2}{4}} &= \frac{\bar{\mathfrak{h}} - \underline{\mathfrak{h}}}{\sqrt{\bar{\mathfrak{h}} + \frac{|e(t)|^2}{4}} + \sqrt{\underline{\mathfrak{h}} + \frac{|e(t)|^2}{4}}}, \\ &\geq \frac{\bar{\mathfrak{h}} - \underline{\mathfrak{h}}}{2\sqrt{1 - |e(t)| + \frac{|e(t)|^2}{4}}} \geq \frac{\bar{\mathfrak{h}} - \underline{\mathfrak{h}}}{2}. \end{aligned}$$

When $|e(t)| = 0$, this maximization problem reduces to :

$$\bar{\eta} = \sup_{z: \underline{\mathfrak{h}} \leq |z|^2 \leq \bar{\mathfrak{h}}} \{|z|^4 - |z|^2\}.$$

Since $\bar{\mathfrak{h}} \geq \underline{\mathfrak{h}} = \frac{3}{4}$, its solution is :

$$\bar{\eta} = \bar{\mathfrak{h}}^2 - \bar{\mathfrak{h}}.$$

When $|e(t)| \neq 0$, according to the Karush-Kuhn-Tucker conditions, a necessary condition for z to be a maximizer is the existence of real numbers $\underline{\lambda}$ and $\bar{\lambda}$ satisfying :

$$\begin{aligned} \frac{\partial}{\partial z} \left\{ \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 + [\underline{\lambda} + \bar{\lambda}] \left[|z|^2 - e(t)^\top z \right] \right\} &= 0 \\ \min \left\{ \underline{\lambda}, |z|^2 - e(t)^\top z - \underline{\mathfrak{h}} \right\} &= 0, \quad \max \left\{ \bar{\lambda}, |z|^2 - e(t)^\top z - \bar{\mathfrak{h}} \right\} = 0. \end{aligned}$$

Hence, when $\bar{\lambda} \neq 0$, respectively $\underline{\lambda} \neq 0$, the maximizer z_{sup} is located at a point of tangency of the two planar curves :

$$\begin{aligned} \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 &= \eta, \\ |z|^2 - e(t)^\top z &= \bar{\mathfrak{h}}, \quad \text{respectively} = \underline{\mathfrak{h}}, \end{aligned}$$

for some η . So the gradients must be colinear, i.e.

$$0 = \left((|z|^2 - e(t)^\top z)(2z - e(t)) - z \right) \wedge (2z - e(t)) = z \wedge e(t).$$

This implies the maximizer z_{sup} is colinear with $e(t)$ and there exists a real number ζ solution of :

$$z_{\text{sup}} = \zeta e(t), \quad (\zeta^2 - \zeta)|e(t)|^2 = \bar{\mathfrak{h}}, \quad \text{respectively} = \underline{\mathfrak{h}}.$$

We get :

$$\zeta = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{\bar{\mathfrak{h}}}{|e(t)|^2}}, \quad \text{respectively} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{\underline{\mathfrak{h}}}{|e(t)|^2}}$$

and

$$\eta = \bar{\mathfrak{h}}^2 - \left[\frac{|e(t)|}{2} - \sqrt{\frac{|e(t)|^2}{4} + \bar{\mathfrak{h}}} \right]^2 = \bar{\eta}(\bar{\mathfrak{h}}), \quad \text{respectively} = \underline{\mathfrak{h}}^2 - \left[\frac{|e(t)|}{2} - \sqrt{\frac{|e(t)|^2}{4} + \underline{\mathfrak{h}}} \right]^2$$

Since, according to Lemma 1, the function $\bar{\mathfrak{h}} \mapsto \bar{\eta}(\bar{\mathfrak{h}})$ defined in (23) is strictly increasing on $\left] \frac{1}{2}, +\infty \right[$, the largest value for η is :

$$\eta = \bar{\mathfrak{h}}^2 - \left[\frac{|e(t)|}{2} - \sqrt{\frac{|e(t)|^2}{4} + \bar{\mathfrak{h}}} \right]^2.$$

This formula covers also the case $|e(t)| = 0$.

When $\bar{\lambda} = \underline{\lambda} = 0$, the maximizer z_{sup} is a zero of the gradient of \mathcal{C}_z and therefore \bar{z} , \underline{z} or \underline{z} . But \bar{z} and \underline{z} do not satisfy

$$|z|^2 - e(t)^\top z \geq \underline{\mathfrak{h}} .$$

Also, although the global minimizer \underline{z} of \mathcal{C}_z may satisfy this inequality for some $|e(t)|$, its corresponding value :

$$\eta = \left(|\underline{z}|^2 - e(t)^\top \underline{z} \right)^2 - |\underline{z}|^2$$

is smaller than all the possible ones.

All this implies that $\bar{\eta}(\bar{\mathfrak{h}})$ in (47) is

$$\begin{aligned} \bar{\eta}(\bar{\mathfrak{h}}) &= \left(|z_{\text{sup}}|^2 - e(t)^\top z_{\text{sup}} \right)^2 - |z_{\text{sup}}|^2 \Big|_{z_{\text{sup}} = \left(\frac{|e(t)|}{2} - \sqrt{\frac{|e(t)|^2}{4} + \bar{\mathfrak{h}}} \right) \frac{e(t)}{|e(t)|}} , \\ &= \bar{\mathfrak{h}}^2 - \left[\frac{|e(t)|}{2} - \sqrt{\frac{|e(t)|^2}{4} + \bar{\mathfrak{h}}} \right]^2 . \end{aligned}$$

In conclusion, we have established that, for all $\bar{\mathfrak{h}}$ in $\left[\frac{3-2|e(t)|}{4}, 1-|e(t)| \right]$, all (z, t) in the super level :

$$\left\{ (z, t) \in \mathcal{D} : \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 \geq \bar{\eta}(\bar{\mathfrak{h}}) \right\} ,$$

satisfy :

$$\frac{\partial^2 \mathcal{C}_z}{\partial z \partial z}(z, t) \geq 2\wp \quad , \quad d \left(z, \left\{ \zeta \in \mathbb{R}^2 : |\zeta|^2 - e(t)^\top \zeta \leq \frac{3-2|e(t)|}{4} \right\} \right) \geq \frac{\bar{\mathfrak{h}} - \frac{3-2|e(t)|}{4}}{2} .$$

B.4 About point 4 in Proposition 4

We want to show that, for each t , each z_a and z_b such that (z_a, t) and (z_b, t) are in the super level set

$$\Omega^{3\varepsilon} = \left\{ (z, t) \in \mathcal{D} : \left(|z|^2 - e(t)^\top z \right)^2 - |z|^2 \geq \bar{\eta} \left(\frac{3-2|e(t)|}{4} \right) \right\} .$$

and for all s in $[0, 1]$ such that, with :

$$z_s = sz_a + (1-s)z_b ,$$

(z_s, t) is also in $\Omega^{3\varepsilon}$, we have :

$$s\mathcal{C}_z(z_a, t) + (1-s)\mathcal{C}_z(z_b, t) \geq \mathcal{C}_z(z_s, t) .$$

Actually this inequality is a straightforward consequence of the following :

$$(i) \quad \mathcal{C}_z(z_a, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_a - z_s) ,$$

$$(ii) \quad \mathcal{C}_z(z_b, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_b - z_s) ,$$

As in (9), point 3 of Proposition 4 implies :

$$\mathcal{C}_z(z_\beta, t) \geq \mathcal{C}_z(z_\alpha, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_\alpha, t)(z_\beta - z_\alpha) + \wp |z_\alpha - z_\beta|^2 \quad (48)$$

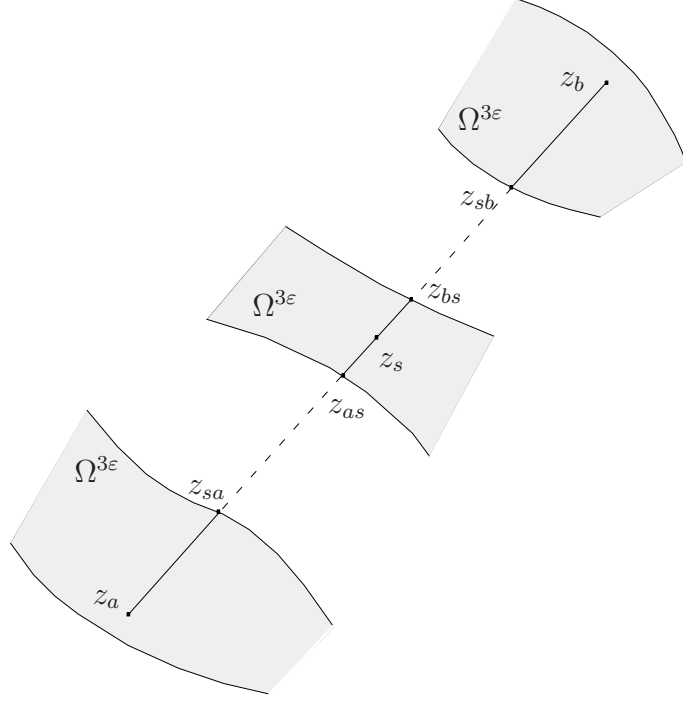


Figure 4: Illustration of the set $\Omega^{3\varepsilon}$.

provided the segment $[z_\alpha, z_\beta]$ is entirely contained in $\Omega^{3\varepsilon}$. So we are done if the entire segment $[z_a, z_b]$ is contained in $\Omega^{3\varepsilon}$.

If it is not, by assumption, z_s is in $\Omega^{3\varepsilon}$. In this case we introduce the notations z_{sa} , z_{as} , z_{bs} and z_{sb} , reported in Figure 4. These points are defined to satisfy :

$$\mathcal{C}_z(z_{sa}, t) = \mathcal{C}_z(z_{as}, t) = \mathcal{C}_z(z_{bs}, t) = \mathcal{C}_z(z_{sb}, t) = \bar{\eta} \left(\frac{3 - 2|e(t)|}{4} \right)$$

and the 3 segments $[z_a, z_{sa}]$, $[z_{as}, z_{bs}]$ and $[z_{sb}, z_b]$ are each entirely contained in $\Omega^{3\varepsilon}$ with z_s in $[z_{as}, z_{bs}]$. Depending on the configuration, we may have for example $z_{sa} = z_{as}$, in which case, we consider the segment $[z_a, z_{bs}]$, and similarly. Then (48) gives us :

$$\mathcal{C}_z(z_a, t) \geq \mathcal{C}_z(z_{sa}, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_{sa}, t)(z_a - z_{sa}) ,$$

$$\mathcal{C}_z(z_{sa}, t) = \mathcal{C}_z(z_{as}, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_s, t)(z_{as} - z_s) , \quad (49)$$

$$\mathcal{C}_z(z_{sb}, t) = \mathcal{C}_z(z_{bs}, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_s, t)(z_{bs} - z_s) , \quad (50)$$

$$\mathcal{C}_z(z_b, t) \geq \mathcal{C}_z(z_{sb}, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_{sb}, t)(z_b - z_{sb}) .$$

This yields :

$$\mathcal{C}_z(z_a, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_s, t)(z_{as} - z_s) + \frac{\partial \mathcal{C}_z}{\partial x}(z_{sa}, t)(z_a - z_{sa}) , \quad (51)$$

$$\mathcal{C}_z(z_b, t) \geq \mathcal{C}_z(z_s, t) + \frac{\partial \mathcal{C}_z}{\partial x}(z_s, t)(z_{bs} - z_s) + \frac{\partial \mathcal{C}_z}{\partial x}(z_{sb}, t)(z_b - z_{sb}) .$$

By definition, any point in the segment $[z_a, z_{sa}]$, respectively $[z_{sb}, z_b]$ is in the super level set $\Omega^{3\varepsilon}$. This implies :

$$\mathcal{C}_z(z_{sa} + r[z_a - z_{sa}]) \geq \mathcal{C}_z(z_{sa}, t) \quad , \quad \mathcal{C}_z(z_{sb}, t) \leq \mathcal{C}_z((z_{sb} + r[z_b - z_{sb}]), t) \quad \forall r \in [0, 1]$$

and, consequently :

$$\frac{\partial \mathcal{C}_z}{\partial z}(z_{sa}, t)(z_a - z_{sa}) \geq 0 \quad , \quad \frac{\partial \mathcal{C}_z}{\partial z}(z_{sb}, t)(z_b - z_{sb}) \geq 0 . \quad (52)$$

On another hand, in (49) and (50), we have :

$$0 \geq \mathcal{C}_z(z_{as}, t) - \mathcal{C}_z(z_s, t) \quad , \quad 0 \geq \mathcal{C}_z(z_{bs}, t) - \mathcal{C}_z(z_s, t) .$$

This gives :

$$0 \geq \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_{as} - z_s) \quad , \quad 0 \geq \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_{bs} - z_s) .$$

Now since z_{as} and z_{sa} , respectively z_{bs} and z_{sb} are in the segment $[z_a, z_s]$, respectively $[z_b, z_s]$, we have also :

$$\frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_{as} - z_s) \geq \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_a - z_s) \quad , \quad \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_{bs} - z_s) \geq \frac{\partial \mathcal{C}_z}{\partial z}(z_s, t)(z_b - z_s) . \quad (53)$$

(51), (52) and (53) give (i) and (ii).

C Another way of choosing the output function

We start by noting that solutions of :

$$\mathcal{C}_z(z, t) = (|z|^2 - e(t)^\top z)^2 - |z|^2 = 0$$

are solutions of :

$$|z|^2 - |z| = e(t)^\top z \quad , \quad |z|^2 + |z| = e(t)^\top z \quad (\leq |e(t)||z|) .$$

When $|e(t)| \leq \frac{1}{2}$, the second equation has no solution. So we have :

$$\left\{ (z, t) : z \neq 0 , \quad \mathcal{C}_z(z, t) = 0 \right\} = \left\{ (z, t) : z \neq 0 , \quad |z|^2 - |z| - e(t)^\top z = 0 \right\} .$$

Actually this level set is also :

$$\left\{ (z, t) : |z| - 1 - e(t)^\top \frac{z}{|z|} = 0 \right\} .$$

The corresponding sub-level set :

$$\left\{ (z, t) : |z| - 1 - e(t)^\top \frac{z}{|z|} \leq 0 \right\} \quad (54)$$

is convex.

Then, to build up h_z , we introduce a t -dependent strictly Hurwitz matrix $A(t)$ and a point $z_e(t)$ in the set $\left\{ z : |z| - 1 - e(t)^\top \frac{z}{|z|} < 0 \right\}$ satisfying :

$$\left[\frac{|z|^2(z - e(t)) + (e(t)^\top z)z}{|z|^3} \right]^\top A(t)[z(s) - z_e(t)] = \frac{\partial}{\partial z} \left\{ |z| - 1 - e(t)^\top \frac{z}{|z|} \right\} A(t)[z(s) - z_e(t)] < 0$$

$$\forall (z, t) : |z| - 1 - e(t)^\top \frac{z}{|z|} = 0 .$$

In this case, if $Z(z, s; t)$ denotes the solution of :

$$\frac{dz}{ds}(s) = A(t)[z(s) - z_e(t)]$$

going through z when $s = 0$, i.e.

$$Z(z, s; t) = \exp(A(t)s)[z - z_e(t)] + z_e(t)$$

Then, for all $(z, t) \neq (z_e(t), t)$, there exists $S(z, t)$ solution of :

$$|Z(z, S(z, t); t)| - 1 - e(t)^\top \frac{Z(z, S(z, t); t)}{|Z(z, S(z, t); t)|} = 0$$

satisfying :

$$\lim_{z \rightarrow z_e(t)} S(z, t) = -\infty .$$

Then we can pick :

$$h_z(z, t) = \exp(S(z, t)) - 1 .$$

The convexity of the sub-level sets of h_z follows from the convexity of (54) and the property that the function $z \mapsto Z(z, s; t)$ transforms straight lines into straight lines for all (s, t) .

For example, with :

$$A = -I \quad , \quad z_e = 0 \quad ,$$

we obtain :

$$\begin{aligned} - \left[\frac{|z|^2(z - e(t)) + (e(t)^\top z)z}{|z|^3} \right]^\top z &= -|z| \quad , \\ &\leq -(1 - |e(t)|) \leq -\frac{1}{2} \quad \forall (z, t) : |z| - 1 - e(t)^\top \frac{z}{|z|} = 0 . \end{aligned}$$

Also the corresponding flow

$$Z(z, s; t) = \exp(-s)z$$

transforms straight lines into straight lines. Then $S(z, t)$ is solution of :

$$\exp(-S(z, t))|z| - 1 - e(t)^\top \frac{z}{|z|} = 0 .$$

This yields :

$$h_z(z, t) = \left(\frac{|z|^2}{|z| + e(t)^\top z} \right) - 1 . \tag{55}$$

It remains to modify this function on a neighborhood of the origin to make it C^2 and meet Assumptions A1, A2 and A3. This can be done with arguments similar to those used in Sections 3.2 and 3.3. Its level sets are shown on Figure 5.

With the notation :

$$tmp = \varphi - (L_0 - L_1)i(t) \quad ,$$

this leads to the flux observer :

$$\dot{\hat{\varphi}} = v - Ri \quad \text{if} \quad \frac{|tmp|^2}{\Phi |tmp| + 2L_1 i(t)^\top tmp} \leq 1$$

and :

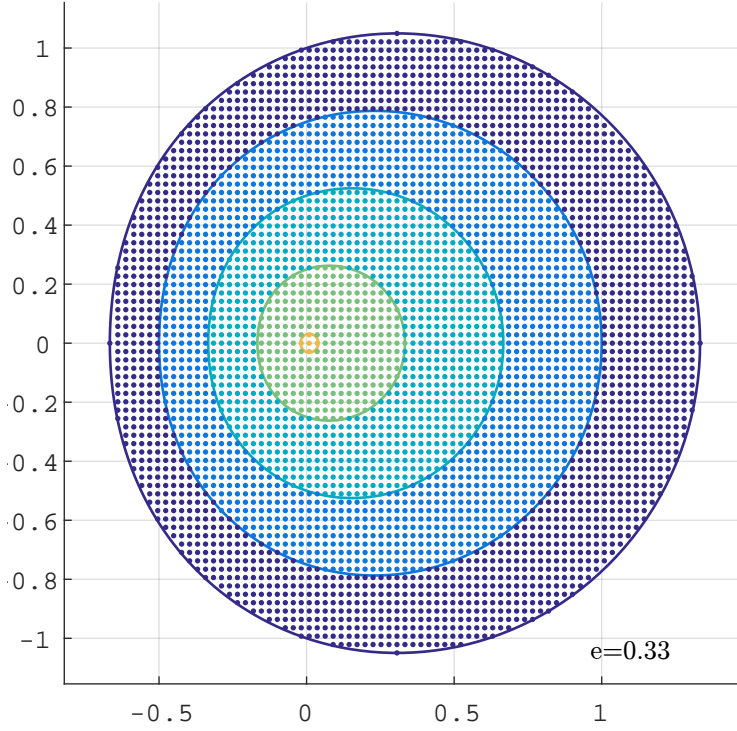


Figure 5: Level sets of h_z in (55)

$$\dot{\hat{\varphi}} = v - Ri + \mu \left[\frac{\left(\Phi |tmp| + 4L_1 i(t)^\top tmp \right) tmp - 2 |tmp|^2 L_1 i(t)}{\left(\Phi |tmp| + 2L_1 i(t)^\top tmp \right)^2} \right] \left[\frac{|tmp|^2}{\Phi |tmp| + 2L_1 i(t)^\top tmp} - 1 \right]$$

if not.

References

- [1] J. Chiasson, Modeling and High-Performance Control of Electric Machines. IEEE Press, 2005
- [2] W. Dib, R. Ortega, J. Malaizé. Sensorless control of permanent-magnet synchronous motor in automotive applications: Estimation of the angular position. Proceedings of the 37th Annual Conference on IEEE Industrial Electronics Society, November 2011
- [3] A. Fradkov, Speed gradient scheme and its applications in adaptive control problem. Automation and Remote Control. Vol. 40, pp. 1333-1342, 1979
- [4] A. Fradkov, R. Ortega, G. Bastin, Semi-adaptive control of convexly parameterized systems with applications to temperature regulation of chemical reactors. International Journal of Adaptive Control and Signal Processing. Vol. 15, 4, Jpp. 415-426, June 2001
- [5] A. Koji, A. M. Annaswamy, A. P. Loh, R. Lozano, Adaptive Control of a Class of Nonlinear Systems With Convex/Concave Parametrization, Systems and Control Letters, Issue: 37, 267-274, 1999.

- [6] J. Lee, J. Hong, K. Nam, R. Ortega, L. Praly, A. Astolfi, Sensorless control of surface-mount permanent magnet synchronous motors based on a nonlinear observer. *IEEE Transactions on Power Electronics*, Vol. 25 , N. 2, February 2010 , Pages 290 - 297
- [7] E. H. Lockwood, *A Book of Curves*. Cambridge University Press, 1961.
See also

<http://www.mathcurve.com/courbes2d.gb/limacon/limacon.shtml>
<https://mathworld.wolfram.com/Limacon.html>
- [8] J. Malaizé, L. Praly, N. Henwood Globally convergent nonlinear observer for the sensorless control of surface-mount permanent magnet synchronous machines. *Proceedings of the 51st IEEE Conference on Decision and Control*, December 2012.
- [9] R. Ortega, L. Praly, A. Astolfi, J. Lee, K. Nam, Estimation of Rotor Position and Speed of Permanent Magnet Synchronous Motors With Guaranteed Stability. *IEEE Transactions on Control Systems Technology*, Issue: 99, 1 - 13, 2010
- [10] F. Poulain, L. Praly, R. Ortega. An Observer for Permanent Magnet Synchronous Motors with Currents and Voltages as only Measurements. *Proceedings of the 47th IEEE Conference on Decision and Control*, December 2008
- [11] R. Sanfelice, L. Praly. Convergence of nonlinear observers on \mathbb{R}^n with a Riemannian metric (Part I). *IEEE Transactions on Automatic Control*, vol. 57, no. 7, july 2012
- [12] B. Sturmfels, *Solving systems of polynomial equations*, American Mathematical Society, 2002
- [13] P. Vas, *Sensorless Vector and Direct Torque Control*. Oxford Science Publication, 1998
- [14] P. Vas, *Electrical machines and drives: a space-vector theory approach*, Oxford University Press, 1992
- [15] M. Yan, Extension of convex functions, *Journal of Convex Analysis* 21 (2014), No. 4, pp. 965-987