Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors

Philippe Martin, Erwan Salaün

Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007

Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors

Philippe Martin and Erwan Salaün

Abstract— **In this paper we propose invariant nonlinear observers for attitude and heading estimation using directly the measurements from low-cost inertial and magnetic sensors. In particular we propose a simple and easy-to-tune observer where moreover the estimated attitude behaves well even in the presence of magnetic disturbances.**

I. INTRODUCTION

Aircraft, especially Unmanned Aerial Vehicles (UAV), usually need to know their orientation to be operated, whether manually or with computer assistance. When cost or weight is an issue, using very accurate inertial sensors for "true" (i.e. based on the Schuler effect due to a non-flat rotating Earth) inertial navigation is excluded. Instead, low-cost systems for attitude and heading estimation –often called Attitude Heading Reference Systems (AHRS)– rely on light and cheap "stapdown" gyroscopes, accelerometers and magnetometers; they "merge" the various measurements according to the motion equations of the aircraft, assuming a flat non-rotating Earth and a low acceleration, usually with a linear complementary filter or an Extended Kalman Filter (EKF). For more details about avionics and various inertial navigation systems, see for instance [2], [4].

Recently, several nonlinear observers have also been suggested. The most interesting is [3], since it uses directly the measurements, whereas [7], [5] must first algebraically compute the orientation, which debases the performance and increases the computational load.

Pursuing along these lines, we follow in this paper the general method developed in [1] to derive three invariant nonlinear observers using directly the measurements. These observers have by design a nice geometrical structure which respects natural symmetries, are simple and easy to tune, and yield (at least local) convergence around every trajectory of the system. Moreover they are computationally much simpler than an EKF. We illustrate their performances on experimental data and study their behavior in the presence of magnetic disturbances.

The authors are with the Centre Automatique et Systèmes, École des Mines de Paris, 60 boulevard Saint-Michel, 75272 Paris Cedex 06, FRANCE philippe.martin@ensmp.fr, erwan.salaun@ensmp.fr

II. PHYSICAL EQUATIONS AND MEASUREMENTS

A. Motion equations

The motion of a flying rigid body (assuming the Earth is flat and defines an inertial frame) is described by

$$\dot{q} = \frac{1}{2} q * \omega$$
$$\dot{V} = A + q * a * q^{-1},$$

where

- q is the quaternion representing the orientation of the body-fixed frame with respect to the Earth-fixed frame
- ω is the instantaneous angular velocity vector
- V is the velocity vector of the center of mass with respect to the Earth-fixed frame
- A is the (constant) gravity vector
- a is the specific acceleration vector, and represents here the aerodynamics forces divided by the body mass.

The first equation describes the kinematics of the body, the second is Newton's force law. It is customary to use quaternions (also called Euler 4-parameters) instead of Euler angles since they provide a global parametrization of the body orientation, and are well-suited for calculations and computer simulations. For more details about this section, see any good textbook on aircraft modelling, for instance [6], and section VIII for useful formulas used in this paper.

B. Measurements

We have three triaxial sensors, providing nine scalar measurements. The gyros measure $\omega_m = \omega + \omega_b$, where the bias ω_b is supposed to be a constant vector. The magnetometers measure $y_b = q^{-1} * B * q$, where $B = (B_1\ 0\ B_3)^T$ is the Earth magnetic field in North-East-Down coordinates. The hypothesis of low acceleration allows us to consider $\dot{V} = 0$. Therefore we can approximate the accelerometers measurements by $a = -q^{-1} * A * q$, where $A = (0\ 0\ g)^T$ in NED coordinates.

C. The considered system

We thus consider the following system:

$$\dot{q} = \frac{1}{2} q * (\omega_m - \omega_b) \tag{1}$$
$$\dot{\omega}_b = 0 \tag{2}$$

with the output

$$\begin{pmatrix} y_A \\ y_C \end{pmatrix} = \begin{pmatrix} q^{-1} * A * q \\ q^{-1} * C * q \end{pmatrix} \tag{3}$$

where $C = B$ or $C = A \times B$ since we can also use $y_A \times y_B = y_{A\times B}$. These are the dynamic equations used to construct our observers, whose inputs are the three gyroscopes measurements ω_m.

III. THEORY OF INVARIANT OBSERVERS

We briefly recall here the main ideas of [1].

A. Invariant systems and compatible outputs

Definition 1: Let G be a Lie Group with identity e and Σ an open set (or more generally a manifold). A *transformation group* $(\phi_g)_{g\in G}$ on Σ is a smooth map

$$(g,\xi) \in G\times\Sigma \mapsto \phi_g(\xi) \in \Sigma$$

such that:

- $\phi_e(\xi) = \xi$ for all ξ
- $\phi_{g_2}\circ\phi_{g_1}(\xi) = \phi_{g_2 g_1}(\xi)$ for all g_1, g_2, ξ.

By construction ϕ_g is a diffeomorphism on Σ for all g. The transformation group is *local* if $\phi_g(\xi)$ is defined only for g around e. In this case the transformation law $\phi_{g_2}\circ\phi_{g_1}(\xi) = \phi_{g_2 g_1}(\xi)$ is imposed only when it makes sense. We consider in the sequel only local transformation groups. "For all g" thus means "for all g around e, and "for all ξ" means "for all ξ in some neighborhood".

Consider now the smooth output system

$$\dot{x} = f(x,u) \tag{4}$$

$$y = h(x,u) \tag{5}$$

where x belongs to an open subset $\mathscr{X} \subset \mathbb{R}^n$, u to an open subset $\mathscr{U} \subset \mathbb{R}^m$ and y to an open subset $\mathscr{Y} \subset \mathbb{R}^p$, $p \le n$.

We assume the signals $u(t), y(t)$ known (y is measured, and u is measured or known as a control input).

Consider also the local group of transformations on $\mathscr{X}\times\mathscr{U}$ defined by $(X,U) = \big(\varphi_g(x), \psi_g(u)\big)$, where φ_g and ψ_g are local diffeomorphisms.

Definition 2: The system $\dot{x} = f(x,u)$ is *G-invariant* if $f\big(\varphi_g(x),\psi_g(u)\big) = D\varphi_g(x)\cdot f(x,u)$ for all g,x,u.
The property also reads $\dot{X} = f(X,U)$, i.e., the system is left unchanged by the transformation.

Definition 3: The output $y = h(x,u)$ is *G-compatible* if there exists a transformation group $(\rho_g)_{g\in G}$ on $\mathscr{Y}$ such that $h\big(\varphi_g(x),\psi_g(u)\big) = \rho_g\big(h(x,u)\big)$ for all g,x,u.
With $(X,U) = \big(\varphi_g(x),\psi_g(u)\big)$ and $Y = \rho_g(y)$, the definition reads $Y = h(X,U)$.

B. Invariant pre-observers

Definition 4 (Pre-observer): The system $\dot{\hat{x}} = F(\hat{x},u,y)$ is a *pre-observer* of (4)-(5) if $F\big(x,u,h(x)\big) = f(x,u)$ for all x,u. An observer is then a pre-observer such that $\hat{x}(t) \to x(t)$ (possibly only locally).

Definition 5: The pre-observer $\dot{\hat{x}} = F(\hat{x},u,y)$ is *G-invariant* if $F\big(\varphi_g(\hat{x}),\psi_g(u),\rho_g(y)\big) = D\varphi_g(\hat{x})\cdot F(\hat{x},u,y)$ for all $g,\hat{x},u,y$.
The property also reads $\dot{\hat{X}} = F(\hat{X},U,Y)$, i.e., the system is left unchanged by the transformation.

The key idea to build an invariant observer is to use an invariant output error.

Definition 6: The smooth map $(\hat{x},u,y) \mapsto E(\hat{x},u,y) \in \mathscr{Y}$ is an *invariant output error* if

- the map $y \mapsto E(\hat{x},u,y)$ is invertible for all $\hat{x},u$
- $E\big(\hat{x},u,h(\hat{x},u)\big) = 0$ for all $\hat{x},u$
- $E\big(\varphi_g(\hat{x}),\psi_g(u),\rho_g(y)\big) = E(\hat{x},u,y)$ for all $\hat{x},u,y$

The first and second properties mean E is an "output error", i.e. it is zero if and only if $h(\hat{x},u) = y$; the third property, which also reads $E(\hat{X},U,Y) = E(\hat{x},u,y)$, defines invariance.

Similarly, the key idea to study the convergence of an invariant observer is to use an invariant state error.

Definition 7: The smooth map $(\hat{x},x) \mapsto \eta(\hat{x},x) \in \mathscr{X}$ is an *invariant state error* if

- it is a diffeomorphism on $\mathscr{X}\times\mathscr{X}$
- $\eta(x,x) = 0$ for all x
- $\eta\big(\varphi_g(\hat{x}),\varphi_g(x)\big) = \eta(\hat{x},x)$ for all $\hat{x},x$.

We now state the two main results –based on the Cartan moving frame method– in the special case where $g \mapsto \varphi_g(x)$ is invertible (i.e. when G is of dimension n), see [1] for the general case. The *moving frame* $x \mapsto \gamma(x)$ is obtained by solving for g the so-called *normalization equation* $\varphi_g(x) = c$ for some arbitrary constant c; in other words $\varphi_{\gamma(x)}(x) = c$.

Theorem 1: The general invariant pre-observer reads

$$F(\hat{x},u,y) = f(\hat{x},u) + \sum_{i=1}^{n}\big(L_i(E,I)\cdot E\big)w_i(\hat{x}),$$

where:

- w_i, $i = 1,\dots,n$, is the invariant vector field defined by
$$w_i(\hat{x}) = \big[D\varphi_{\gamma(\hat{x})}(\hat{x})\big]^{-1}\cdot\frac{\partial}{\partial x_i},$$
with $\frac{\partial}{\partial x_i}$ the i^{th} canonical vector field on $\mathscr{X}$
- E is the invariant error defined by
$$E(\hat{x},u,y) = \rho_{\gamma(\hat{x})}\big(h(\hat{x},u)\big) - \rho_{\gamma(\hat{x})}(y)$$
- I is the (complete) invariant defined by
$$I(\hat{x},u) = \psi_{\gamma(\hat{x})}(u)$$
- L_i, $i = 1,\dots,n$, is a $1\times p$ matrix with entries possibly depending on E and I, and can be freely chosen.

Theorem 2: The error system reads $\dot{\eta} = \Upsilon(\eta, I)$ where η is the invariant state error defined by

$$\eta(\hat{x},x) = \varphi_{\gamma(x)}(\hat{x}) - \varphi_{\gamma(x)}(x).$$

This result greatly simplifies the convergence analysis of the pre-observer, since the error equation is autonomous but for the "free" known invariant I. For a general nonlinear (not invariant) observer the error equation depends on the trajectory $t \mapsto \big(x(t),u(t)\big)$ of the system, hence is in fact of dimension $2n$.

IV. OBSERVERS INVARIANT BY BODY-FIXED ROTATIONS AND TRANSLATIONS

A. Invariance of the system equations

All the measurements are expressed in the body-fixed frame. From a physical and engineering viewpoint, a sensible observer using these measurements should not be affected by the actual choice of body-fixed coordinates, i.e. by a constant rotation in the body-fixed frame. Similarly, a translation of the gyro bias by a vector constant in the body-fixed frame should not affect the observer.

We therefore consider the transformation group generated by constant rotations and translations in the body-fixed frame

$$\varphi_{(q_0,\omega_0)}\begin{pmatrix} q \\ \omega_b \end{pmatrix} = \begin{pmatrix} q*q_0 \\ q_0^{-1}*\omega_b*q_0+\omega_0 \end{pmatrix}$$
$$\psi_{(q_0,\omega_0)}(\omega_m) = q_0^{-1}*\omega_m*q_0+\omega_0$$
$$\rho_{(q_0,\omega_0)}\begin{pmatrix} y_A \\ y_C \end{pmatrix} = \begin{pmatrix} q_0^{-1}*y_A*q_0 \\ q_0^{-1}*y_C*q_0 \end{pmatrix},$$

where q_0 is a unit quaternion and ω_0 a vector in $\mathbb{R}^3$. It is indeed a transformation group since

$$\varphi_{(q_1,\omega_1)}\circ\varphi_{(q_0,\omega_0)}\begin{pmatrix} q \\ \omega_b \end{pmatrix} = \varphi_{(q_1,\omega_1)\diamond(q_0,\omega_0)}\begin{pmatrix} q \\ \omega_b \end{pmatrix}$$
$$\psi_{(q_1,\omega_1)}\circ\psi_{(q_0,\omega_0)}(\omega_m) = \psi_{(q_1,\omega_1)\diamond(q_0,\omega_0)}(\omega_m)$$
$$\rho_{(q_1,\omega_1)}\circ\rho_{(q_0,\omega_0)}\begin{pmatrix} y_A \\ y_C \end{pmatrix} = \rho_{(q_1,\omega_1)\diamond(q_0,\omega_0)}\begin{pmatrix} y_A \\ y_C \end{pmatrix},$$

where the group composition law $\diamond$ is defined by

$$(q_1,\omega_1)\diamond(q_0,\omega_0) = (q_0*q_1, q_1^{-1}\omega_0 q_1+\omega_1).$$

The system (1)-(2) is of course invariant by the transformation group since

$$\begin{aligned}
\dot{\overbrace{(q*q_0)}} &= \dot{q}*q_0 \\
&= \frac{1}{2}(q*q_0)*\big((q_0^{-1}*\omega_m*q_0+\omega_0) \\
&\quad -(q_0^{-1}*\omega_b*q_0+\omega_0)\big) \\
\dot{\overbrace{(q_0^{-1}*\omega_b*q_0+\omega_0)}} &= q_0^{-1}*\dot{\omega}_b*q_0 \\
&= 0,
\end{aligned}$$

whereas the output (3) is compatible since

$$\begin{pmatrix} (q*q_0)^{-1}*A*(q*q_0) \\ (q*q_0)^{-1}*C*(q*q_0) \end{pmatrix} = \rho_{(q_0,\omega_0)}\begin{pmatrix} q^{-1}*A*q \\ q^{-1}*C*q \end{pmatrix}.$$

B. Construction of the general invariant pre-observer

We solve for (q_0,ω_0) the normalization equations

$$q*q_0 = 1$$
$$q_0^{-1}*\omega_b*q_0+\omega_0 = 0$$

to find the moving frame

$$\gamma(q,\omega_b) = \begin{pmatrix} q^{-1} \\ -q*\omega_b*q^{-1} \end{pmatrix}.$$

We then get the 6-dimensional invariant error

$$\begin{aligned}
\begin{pmatrix} E_A(\hat{q},\hat{\omega}_b,y_A) \\ E_C(\hat{q},\hat{\omega}_b,y_C) \end{pmatrix} &= \rho_{\gamma(\hat{q},\hat{\omega}_b)}\begin{pmatrix} \hat{q}^{-1}*A*\hat{q} \\ \hat{q}^{-1}*C*\hat{q} \end{pmatrix} - \rho_{\gamma(\hat{q},\hat{\omega}_b)}\begin{pmatrix} y_A \\ y_C \end{pmatrix} \\
&= \begin{pmatrix} A-\hat{q}*y_A*\hat{q}^{-1} \\ C-\hat{q}*y_C*\hat{q}^{-1} \end{pmatrix}
\end{aligned}$$

and the 3-dimensional complete invariant

$$I(\hat{q},\hat{\omega}_b,\omega_m) = \psi_{\gamma(\hat{q},\hat{\omega}_b)}(\omega_m) = \hat{q}*(\omega_m-\hat{\omega}_b)*\hat{q}^{-1}.$$

It is straightforward to check that E_A, E_C and I are indeed invariant. For instance,

$$\begin{aligned}
&E_A(\hat{q}*q_0, q_0^{-1}*\hat{\omega}_b*q_0+\omega_0, q_0^{-1}*y_A*q_0) \\
&\qquad = A-(\hat{q}*q_0)*(q_0^{-1}*y_A*q_0))*(\hat{q}*q_0)^{-1} \\
&\qquad = A-\hat{q}*y_A*\hat{q}^{-1} \\
&\qquad = E_A(\hat{q},\hat{\omega}_b,y_A).
\end{aligned}$$

To find invariant vector fields, we solve for $w(q,\omega_b)$ the 6 vector equations

$$\left[D\varphi_{\gamma(q,\omega_b)}\begin{pmatrix} q \\ \omega_b \end{pmatrix}\right]\cdot w(q,\omega_b) = \begin{pmatrix} e_i \\ 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 \\ e_i \end{pmatrix}, \quad i=1,2,3,$$

where the e_i's are the canonical basis of $\mathbb{R}^3$ (we have identified the tangent space of the unit norm quaternions space to $\mathbb{R}^3$). Since

$$\left[D\varphi_{(q_0,\omega_0)}\begin{pmatrix} q \\ \omega_b \end{pmatrix}\right]\cdot\begin{pmatrix} \delta q \\ \delta\omega_b \end{pmatrix} = \begin{pmatrix} \delta q*q_0 \\ q_0^{-1}*\delta\omega_b*q_0 \end{pmatrix},$$

this yields the 6 independent invariant vector fields

$$\begin{pmatrix} e_i*q \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ q^{-1}*e_i*q \end{pmatrix}, \quad i=1,2,3.$$

It is easy to check that these vector fields are indeed invariant. For instance,

$$\begin{aligned}
\left[D\varphi_{(q_0,\omega_0)}\begin{pmatrix} q \\ \omega_b \end{pmatrix}\right]\cdot\begin{pmatrix} e_i*q \\ 0 \end{pmatrix} &= \begin{pmatrix} (e_i*q)*q_0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} e_i*(q*q_0) \\ 0 \end{pmatrix}.
\end{aligned}$$

The general invariant pre-observer then reads

$$\dot{\hat{q}} = \frac{1}{2}\hat{q}*(\omega_m-\hat{\omega}_b)+\sum_{i=1}^{3}(L_{Ai}E_A+L_{Ci}E_C)e_i*\hat{q}$$
$$\dot{\hat{\omega}}_b = \sum_{i=1}^{3}\hat{q}^{-1}*(M_{Ai}E_A+M_{Ci}E_C)e_i*\hat{q},$$

where the L_{Ai}, L_{Ci}, M_{Ai}, M_{Ci}'s are arbitrary 1×3 matrices with entries possibly depending on E_A, E_C, and I. Noticing

$$\sum_{i=1}^{3}(L_{Ai}E_A)e_i = \begin{pmatrix} L_{A1} \\ L_{A2} \\ L_{A3} \end{pmatrix} E_A = L_A E_A,$$

where L_A is the 3×3 matrix whose rows are the L_{Ai}'s, and defining L_B, M_A and M_B in the same way, we can rewrite the pre-observer as

$$\dot{\hat{q}} = \frac{1}{2}\hat{q}*(\omega_m-\hat{\omega}_b)+(L_AE_A+L_CE_C)*\hat{q} \tag{6}$$
$$\dot{\hat{\omega}}_b = \hat{q}^{-1}*(M_AE_A+M_CE_C)*\hat{q}. \tag{7}$$

As a by-product of its geometric structure, the pre-observer automatically has a desirable feature: the norm of $\hat{q}$ is left unchanged by (6), since $L_AE_A+L_CE_C$ is a vector of $\mathbb{R}^3$ (see sectionVIII).

C. Error equations

The invariant state error is given by

$$\begin{pmatrix}\eta\\ \beta\end{pmatrix} = \varphi_{\gamma(q,\omega_b)}\begin{pmatrix}\hat{q}\\ \hat{\omega}_b\end{pmatrix} - \varphi_{\gamma(q,\omega_b)}\begin{pmatrix}q\\ \omega_b\end{pmatrix} = \begin{pmatrix}\hat{q}*q^{-1}-1\\ q*(\hat{\omega}_b-\omega_b)*q^{-1}\end{pmatrix}.$$

It is in fact more natural –though completely equivalent– to take $\eta = \hat{q}*q^{-1}$ (rather than $\eta = \hat{q}*q^{-1}-1$), so that $\eta(x,x) = 1$, the unit element of the group of quaternions. Hence,

$$\begin{aligned}
\dot{\eta} &= \dot{\hat{q}}*q^{-1} + \hat{q}*(-q^{-1}*\dot{q}*q^{-1})\\
&= -\frac{1}{2}\eta*\beta + (L_AE_A + L_CE_C)*\eta\\
\dot{\beta} &= \dot{q}*(\hat{\omega}_b-\omega_b)*q^{-1} - q*(\hat{\omega}_b-\omega_b)*q^{-1}*\dot{q}*q^{-1}\\
&\quad + q*(\dot{\hat{\omega}}_b - \dot{\omega}_b)*q^{-1}\\
&= (\eta^{-1}*I*\eta)\times\beta + \eta^{-1}*(M_AE_A+M_CE_C)*\eta.
\end{aligned}$$

Since we can write

$$\begin{aligned}
E_A &= A - \eta*A*\eta^{-1}\\
E_C &= C - \eta*C*\eta^{-1},
\end{aligned}$$

we find as expected that the error system

$$\dot{\eta} = -\frac{1}{2}\eta*\beta + (L_AE_A+L_CE_C)*\eta \tag{8}$$

$$\dot{\beta} = (\eta^{-1}*I*\eta)\times\beta + \eta^{-1}*(M_AE_A+M_CE_C)*\eta \tag{9}$$

depends only on the invariant state error (η,β) and the "free" known invariant I, but not on the trajectory of the observed system (1)-(2).

From now on, we assume for simplicity that the gain matrices L_A, L_C, M_A, M_C are constant. We linearize the error system around the equilibrium points $(\bar{\eta},\bar{\beta})$ defined by

$$L_AE_A(\bar{\eta}) + L_CE_C(\bar{\eta}) = M_AE_A(\bar{\eta}) + M_CE_C(\bar{\eta}) = 0, \tag{10}$$

which implies $\bar{\beta} = 0$. We find

$$\begin{aligned}
\delta\dot{\eta} &= -\frac{1}{2}\bar{\eta}*\delta\beta + (L_A\delta E_A + L_C\delta E_C)*\bar{\eta}\\
\delta\dot{\beta} &= (\bar{\eta}^{-1}*I*\bar{\eta})\times\delta\beta + \bar{\eta}^{-1}*(M_A\delta E_A + M_C\delta E_C)*\bar{\eta}
\end{aligned}$$

where

$$\begin{aligned}
\delta E_A &= -\delta\eta*A*\bar{\eta}^{-1} + \bar{\eta}*A*(\bar{\eta}^{-1}*\delta\eta*\bar{\eta}^{-1})\\
&= 2(\bar{\eta}*A*\bar{\eta}^{-1})\times(\delta\eta*\bar{\eta}^{-1})\\
\delta E_C &= 2(\bar{\eta}*C*\bar{\eta}^{-1})\times(\delta\eta*\bar{\eta}^{-1}).
\end{aligned}$$

When moreover $\bar{\eta} = 1$, i.e. the state estimated by the observer equals the actual state, this system reads

$$\begin{aligned}
\delta\dot{\eta} &= -\frac{1}{2}\delta\beta + 2L_A(A\times\delta\eta) + 2L_C(C\times\delta\eta)\\
\delta\dot{\beta} &= I\times\delta\beta + 2M_A(A\times\delta\eta) + 2M_C(C\times\delta\eta).
\end{aligned}$$

V. Choice of the gain matrices L_A, L_C, M_A, M_C

A natural idea is to consider the measurements y_A and y_B on the same footing and to take $C = B$. To ensure the convergence of the linearized error system, a simple choice for L_A, L_C, M_A, M_C is

$$\begin{aligned}
L_AE_A &= l_A A\times E_A \qquad & L_CE_C &= l_C C\times E_C\\
M_AE_A &= -m_A A\times E_A \qquad & M_CE_C &= -m_C C\times E_C
\end{aligned} \tag{11}$$

with $(l_A, m_A, l_C, m_C) > 0$. We could prove the local convergence as we do for another choice of the matrices later in this section. Instead we notice

$$\begin{aligned}
(A\times E_A)*\hat{q} &= \hat{q}*\big(\hat{q}^{-1}*(A\times E_A)*\hat{q}\big)\\
&= \hat{q}*\big((\hat{q}^{-1}*A*\hat{q})\times(\hat{q}^{-1}*E_A*\hat{q})\big)\\
&= \hat{q}*\big(\hat{y}_A\times(y_A-\hat{y}_A)\big)\\
&= \hat{q}*(\hat{y}_A\times y_A).
\end{aligned}$$

Hence the observer (6)-(7) becomes

$$\begin{aligned}
\dot{\hat{q}} &= \frac{1}{2}\hat{q}*(\omega_m - \hat{\omega}_b + l_A\hat{y}_A\times y_A + l_C\hat{y}_C\times y_C)\\
\dot{\hat{\omega}}_b &= -m_A\hat{y}_A\times y_A - m_C\hat{y}_C\times y_C,
\end{aligned}$$

which is, written in quaternion form, the observer proposed in [3]. In that paper, the global convergence is proved.

One drawback of this observer is that it requires the values of B_1 and B_3, which depend on the geographic location. This can be overcome by taking $C = A\times B$ and $y_C = y_A\times y_B$. Another drawback, whether $C = B$ or $C = A\times B$, is that all the error variables are coupled. Hence this observer is more difficult to tune and the attitude is perturbed when the Earth magnetic field is disturbed (see the following section), which is common in practice.

To provide some decoupling of the attitude from the magnetic measurements, we take $C = A\times B$ and

$$\begin{aligned}
L_AE_A &= l_A A\times E_A \qquad & L_CE_C &= \frac{l_C}{g^2}\langle C\times E_C, A\rangle A\\
M_AE_A &= -m_A A\times E_A \qquad & M_CE_C &= -\frac{m_C}{g^2}\langle C\times E_C, A\rangle A.
\end{aligned} \tag{12}$$

The scaling of l_C, m_C by $g^2 = \|A\|^2$ ensures that the correction terms (11)-(12) have the same order of magnitude. Notice we do not need to know the values of B_1 and B_3. The linearized error system now reads

$$\delta\dot{\eta} = \mathrm{D}_\eta\delta\eta - \frac{1}{2}\delta\beta \tag{13}$$

$$\delta\dot{\beta} = \mathrm{D}_\beta\delta\eta + I\times\delta\beta \tag{14}$$

where

$$\mathrm{D}_\eta = \begin{pmatrix} -2g^2l_A & 0 & 0\\ 0 & -2g^2l_A & 0\\ 0 & 0 & -2B_1^2g^2l_C\end{pmatrix}$$

$$\mathrm{D}_\beta = \begin{pmatrix} 2g^2m_A & 0 & 0\\ 0 & 2g^2m_A & 0\\ 0 & 0 & 2B_1^2g^2m_C\end{pmatrix}$$

When $I=0$ (i.e. the system is at rest) the system is completely decoupled

$$\delta\dot{\eta}_i = -2g^2 l_A \delta\eta_i - \frac{1}{2}\delta\beta_i$$
$$\delta\dot{\beta}_i = 2g^2 m_A \delta\eta_i$$

$(i=1,2)$ and

$$\delta\dot{\eta}_3 = -2g^2 B_1^2 l_C \delta\eta_3 - \frac{1}{2}\delta\beta_3$$
$$\delta\dot{\beta}_3 = 2g^2 B_1^2 m_C \delta\eta_3,$$

hence it is very easy to tune. When $I \neq 0$ the equations are slightly coupled by the biases errors $\delta\beta$.

We now prove $(\delta\eta, \delta\beta) \to (0,0)$ whatever $(l_A, m_A, l_C, m_C) > 0$. Indeed, consider the Lyapunov function:

$$V = \frac{g^2 m_A}{2}\delta\eta_1^2 + \frac{g^2 m_A}{2}\delta\eta_2^2 + \frac{g^2 B_1^2 m_C}{2}\delta\eta_3^2 + \frac{1}{8}\|\delta\beta\|^2.$$

Differentiating V and using $\langle \delta\beta, I \times \delta\beta \rangle = 0$, we get:

$$\dot{V} = -2g^2\left(g^2 l_A m_A(\delta\eta_1^2 + \delta\eta_2^2) + g^4 B_1^4 l_C m_C \delta\eta_3^2\right) \leq 0$$

Since V is bounded from below by zero, the preceding inequality implies that $V(\delta\eta(t), \delta\beta(t))$ converges as $t \to \infty$. Since

$$\lim_{t\to\infty}\int_0^t \dot{V}(\delta\eta(\tau), \delta\beta(\tau))d\tau = \lim_{t\to\infty} V(\delta\eta(t), \delta\beta(t)) - V(\delta\eta(0), \delta\beta(0)),$$

we conclude $\lim_{t\to\infty}\int_0^t \dot{V}(\delta\eta(\tau), \delta\beta(\tau))d\tau$ exists and is finite. On the other hand, $\dot{V} \leq 0$ also implies

$$0 \leq V(\delta\eta(t), \delta\beta(t)) \leq V(\delta\eta(0), \delta\beta(0))$$

It follows that $\delta\eta(t)$ and $\delta\beta(t)$ are bounded. The equation (13) implies that $\delta\dot{\eta}(t)$ is bounded too, and finally that $\ddot{V}$ is bounded. Hence $\dot{V}$ is uniformly continuous. And we conclude by Barbalat's lemma as $t \to \infty$:

$$\dot{V} \to 0 \quad \Rightarrow \quad \delta\eta \to 0.$$

Integrating (13), we get

$$\int_0^t \delta\dot{\eta}(\tau)d\tau = \delta\eta(t) - \delta\eta(0) = \int_0^t (D_\eta \delta\eta(\tau) - \frac{1}{2}\delta\beta(\tau))d\tau.$$

Since $\delta\eta(t) \to 0$, it follows

$$\lim_{t\to\infty}\int_0^t (D_\eta \delta\eta(\tau) - \frac{1}{2}\delta\beta(\tau))d\tau = -\delta\eta(0).$$

We assume I is bounded, which is physically sensible. Since $\delta\eta(t)$ and $\delta\beta(t)$ are bounded, $D_\eta\delta\dot{\eta}(t) - \frac{1}{2}\delta\dot{\beta}(t)$ is bounded too. Hence $D_\eta\delta\eta(t) - \frac{1}{2}\delta\beta(t)$ is uniformly continuous. Applying Barbalat's lemma leads to :

$$\lim_{t\to\infty}(D_\eta\delta\eta(t) - \frac{1}{2}\delta\beta(t)) = 0$$

Since $\delta\eta \to 0$, we conclude $\delta\beta \to 0$, which ends the proof.

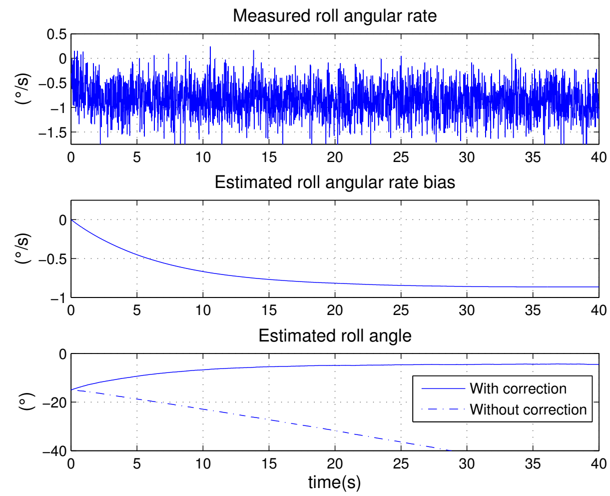

Fig. 1. Motionless estimated roll angle

VI. Experimental Results

The following experimental results have been obtained with the invariant observer

$$\dot{\hat{q}} = \frac{1}{2}\hat{q} * (\omega_m - \hat{\omega}_b) + (L_A E_A + L_C E_C) * \hat{q} - k(1 - \|\hat{q}\|^2)\hat{q}$$
$$\dot{\hat{\omega}}_b = \hat{q}^{-1} * (M_A E_A + M_C E_C) * \hat{q}$$

and the choice of matrices (12). The added term $k(1-\|\hat{q}\|^2)\hat{q}$ is a well-known numerical trick to keep $\|\hat{q}\| = 1$. Notice this term is also invariant.

To have time constants around 30s, we take $l_A = 9e-2$, $m_A = 9e-3$, $l_C = 5e-2$, $m_C = 5e-3$ and $k=1$. We have used for the experimental data the raw measurements of a Microstrain 3DM-GX1. We can compare our results with the estimated orientation given by the 3DM-GX1 (computed according to the user manual by some kind of Kalman filter).

A. *Behavior at rest*

The system is left motionless. We see on Fig. 1 that the bias estimation works well, with a time constant of 30s. This highlights the importance of the correction term in the angle estimation: without correction the estimated roll angle diverges with a slope of $-0.8°/s$ (bottom plot), which is indeed the final value of the estimated bias (middle plot).

B. *Global behavior*

We have not proved yet the global convergence, nevertheless the domain of convergence seems to be big, as can be seen on Fig. 2.

C. *Dynamic behavior*

Once the gyro biases were correctly estimated, we moved the system in all directions. We see on Fig. 3 that the estimated angles are very similar to the angles computed by the 3DM-GX1.

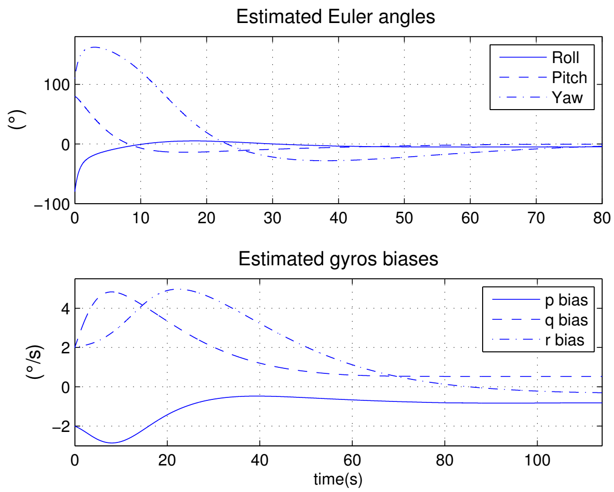

Fig. 2. Motionless estimated gyros biases and Euler angles

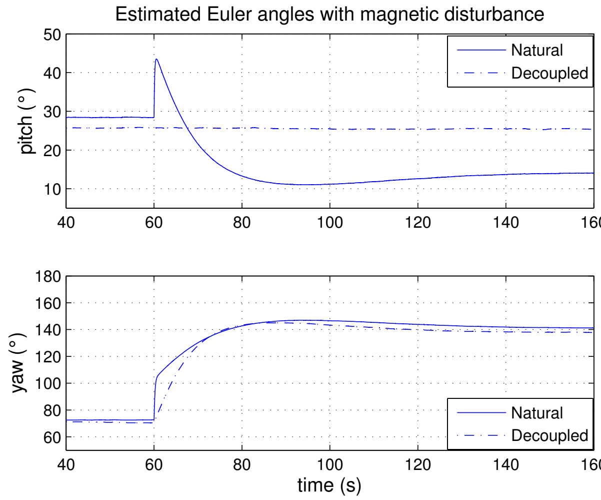

Fig. 4. Estimated pitch and yaw angles with magnetic disturbance

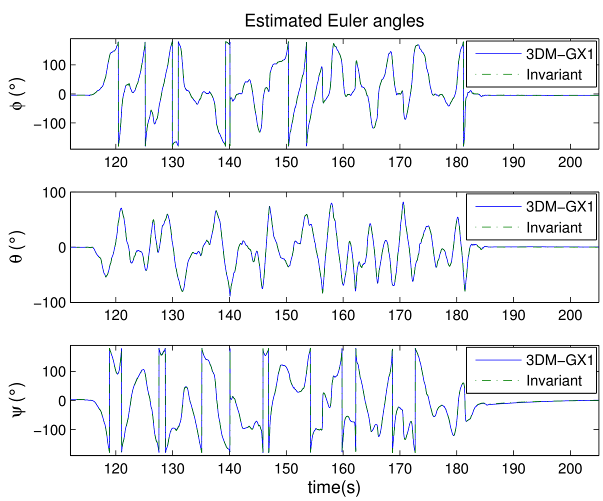

Fig. 3. Dynamic estimated Euler angles

VII. Attitude and Heading Estimation with Magnetic Disturbance

Up to now, we have supposed that the magnetic field measured by the magnetometers was not disturbed. In fact the Earth magnetic field is quite perturbed in an urban area and indoors. If the local magnetic field becomes $B^* = (B_1^*\ B_2^*\ B_3^*)^T$, it affects the observers described above. For aircraft applications, we do not want these disturbances to affect the attitude but only the heading. We now examine the behavior of our observers in the presence of magnetic disturbance. Our analysis requires that the equilibrium points $(\overline{\eta}, \overline{\beta})$ of the error system (8)-(9) satisfies (10). Hence we must assume $L_A E_A + L_C E_C$ colinear to $M_A E_A + M_C E_C$, that is $\frac{m_A}{l_A} = \frac{m_C}{l_C} \triangleq \sigma$. We also denote by $\overline{\phi}, \overline{\theta}, \overline{\psi}$ the Euler angles corresponding to the error quaternion $\overline{\eta}$.

A. *"Natural" observer* $\big(l_A(A \times E_A) + l_C(C \times E_C)$ *with* $C = B\big)$

In this case, $\overline{\phi} = \overline{\theta} = 0$ only if $B_3\sqrt{(B_1^*)^2 + (B_2^*)^2} - B_1 B_3^* = 0$. This is generally not true if $B^* \neq B$. Hence the attitude and not only the heading change because of the magnetic disturbance. The dynamic behavior is of course also modified. See Fig. 4 where a magnet was put near the motionless sensors at $t = 60s$.

B. *Modified "natural" observer* $\big(l_A(A \times E_A) + l_C(C \times E_C)$ *with* $C = A \times B\big)$

In this case, it is easy to check that

$$\overline{\phi} = \overline{\theta} = 0 \quad \text{and} \quad \overline{\psi} = \arctan \frac{C_1^*}{C_2^*}$$

is an equilibrium point which corresponds to the quaternion

$$\overline{\eta} = (\cos\frac{\overline{\psi}}{2}\ \ 0\ \ 0\ \ \sin\frac{\overline{\psi}}{2})^T$$

Only the yaw angle is affected by the magnetic disturbance.

Using the new error variable $\delta\tilde{\eta} = \overline{\eta}^{-1} * \delta\eta$, the linearized error system becomes

$$\begin{aligned}\delta\dot{\tilde{\eta}} = &-\frac{1}{2}\delta\beta + 2l_A\big(A \times (A \times \delta\tilde{\eta})\big)\\ &+ 2l_C\big((\overline{\eta}^{-1} * C * \overline{\eta}) \times (C^* \times \delta\tilde{\eta})\big)\end{aligned}$$

and

$$\begin{aligned}\delta\dot{\beta} =& I \times \delta\beta - 2m_A\big(A \times (A \times \delta\tilde{\eta})\big)\\ &- 2m_C\big((\overline{\eta}^{-1} * C * \overline{\eta}) \times (C^* \times \delta\tilde{\eta})\big)\end{aligned}$$

where we have used $\overline{\eta} * A * \overline{\eta}^{-1} = A$. All the components of $\delta\tilde{\eta}$ are affected by the disturbance. The stability of the equilibrium point can even be changed, for instance if $l_A g + l_C B_1 C_2^* < 0$ and $C_1^* = 0$. Fig. 5 illustrates this case: the equilibrium point $(\overline{\phi}, \overline{\theta}, \overline{\psi}) = (0, 0, \overline{\psi}_1 = 0°)$ before the disturbance is applied becomes unstable and the system moves to the new equilibrium point $(0, 0, \overline{\psi}_2 = 180°)$.

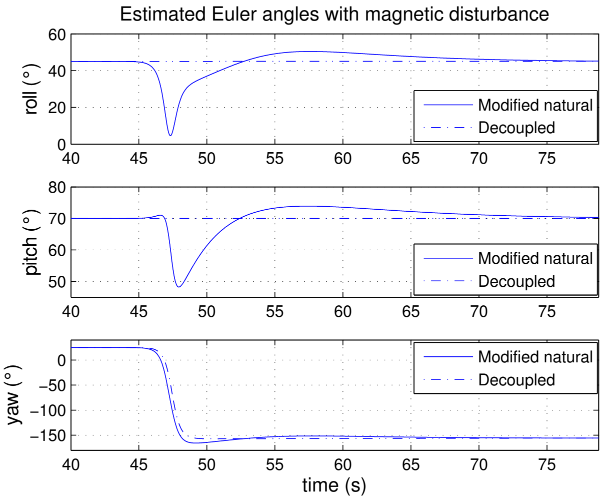

Fig. 5. Estimated Euler angles with magnetic disturbance

C. *Decoupled invariant observer* $\left(l_A(A \times E_A) + \frac{l_C}{g^2}\langle C \times E_C, A\rangle A \text{ with } C = A \times B\right)$

In this case, it is easy to see that as before

$$\overline{\eta} = (\cos\frac{\overline{\psi}}{2}\ 0\ 0\ \sin\frac{\overline{\psi}}{2})^T$$

is an equilibrium point. Only the yaw angle is affected.

The error system for $\delta\tilde{\eta}$ reads

$$\begin{aligned}\delta\dot{\tilde{\eta}} =& -\frac{1}{2}\delta\beta + 2l_A\big(A \times (A \times \delta\tilde{\eta})\big)\\ &+ 2\frac{l_C}{g^2}\langle(\overline{\eta}^{-1} * C * \overline{\eta}) \times (C^* \times \delta\tilde{\eta}), A\rangle A\end{aligned}$$

and

$$\begin{aligned}\delta\dot{\beta} =& I \times \delta\beta - 2m_A\big(A \times (A \times \delta\tilde{\eta})\big)\\ &- 2\frac{m_C}{g^2}\langle(\overline{\eta}^{-1} * C * \overline{\eta}) \times (C^* \times \delta\tilde{\eta}), A\rangle A\end{aligned}$$

where we have used $\overline{\eta} * A * \overline{\eta}^{-1} = A$. We recover the system (13), with the modified matrices

$$\mathrm{D}_\eta = -\frac{1}{\sigma}\mathrm{D}_\beta = \begin{pmatrix} -2g^2 l_A & 0 & 0 \\ 0 & -2g^2 l_A & 0 \\ 0 & 0 & -2B_1 g\|C^*\| l_C \end{pmatrix}.$$

Hence the convergence proof of section V remains the same.

Neither the static nor the dynamic behavior of the attitude is affected by the magnetic disturbance, see Fig. 4 and 5.

VIII. Appendix: quaternions

Thanks to their four coordinates, quaternions provide a global parametrization of the orientation of a rigid body (whereas a parametrization with three Euler angles necessarily has singularities). Indeed, to any quaternion q with unit norm is associated a rotation matrix $R_q \in SO(3)$ by

$$q^{-1} * \vec{p} * q = R_q \cdot \vec{p} \quad \text{for all } \vec{p} \in \mathbb{R}^3.$$

A quaternion p can be thought of as a scalar $p_0 \in \mathbb{R}$ together with a vector $\vec{p} \in \mathbb{R}^3$,

$$p = \begin{pmatrix} p_0 \\ \vec{p} \end{pmatrix}.$$

The (non commutative) quaternion product $*$ then reads

$$p * q \triangleq \begin{pmatrix} p_0 q_0 - \vec{p} \cdot \vec{q} \\ p_0\vec{q} + q_0\vec{p} + \vec{p} \times \vec{q} \end{pmatrix}.$$

The unit element is $e \triangleq \begin{pmatrix} 1 \\ \vec{0} \end{pmatrix}$, and $(p * q)^{-1} = q^{-1} * p^{-1}$.

Any scalar $p_0 \in \mathbb{R}$ can be seen as the quaternion $\begin{pmatrix} p_0 \\ \vec{0} \end{pmatrix}$, and any vector $\vec{p} \in \mathbb{R}^3$ can be seen as the quaternion $\begin{pmatrix} 0 \\ \vec{p} \end{pmatrix}$. We systematically use these identifications in the paper, which greatly simplifies the notations.

We have the useful formulas

$$p \times q \triangleq \vec{p} \times \vec{q} = \frac{1}{2}(p * q - q * p)$$

$$(\vec{p} \cdot \vec{q})\vec{r} = -\frac{1}{2}(p * q + q * p) * r.$$

If q depends on time, then $\dot{q}^{-1} = -q^{-1} * \dot{q} * q^{-1}$.

Finally, consider the differential equation $\dot{q} = q * u + v * q$ where u, v are vectors in $\in \mathbb{R}^3$. Let q^T be defined by $\begin{pmatrix} q_0 \\ -\vec{q} \end{pmatrix}$. Then $q * q^T = \|q\|^2$. Therefore,

$$\overbrace{q * q^T}^{\cdot} = q * (u + u^T) * q^T + \|q\|^2 (v + v^T) = 0$$

since u, v are vectors. Hence the norm of q is constant.

References

[1] S. Bonnabel, Ph. Martin, and P. Rouchon. Invariant observers. *arxiv.math.OC/0612193*, 2007. Submitted to IEEE Trans. Automat. Control.

[2] R.P.G. Collinson. *Introduction to avionics systems*. Kluwer Academic Publishers, second edition, 2003.

[3] T. Hamel and R. Mahony. Attitude estimation on $SO(3)$ based on direct inertial measurements. *Proc. of the 2006 IEEE International Conference on Robotics and Automation*, 2006.

[4] M. Kayton and W.R. Fried, editors. *Avionics navigation systems*. John Wiley & Sons, second edition, 1997.

[5] R. Mahony, T. Hamel, and J-M. Pflimlin. Complementary filter design on the special orthogonal group $SO(3)$. *Proc. of the 44th IEEE Conf. on Decision and Control*, 2005.

[6] B.L. Stevens and F.L. Lewis. *Aircraft control and simulation*. John Wiley & Sons, second edition, 2003.

[7] J. Thienel and R.M. Sanner. A coupled nonlinear spacecraft attitude controller and observer with an unknown constant gyro bias and gyro noise. *IEEE Trans. Automat. Control*, 48(11):2011–2015, 2003.