DESIGN AND IMPLEMENTATION OF A LOW-COST ATTITUDE AND HEADING NONLINEAR ESTIMATOR

Philippe Martin and Erwan Salaün
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

Keywords: Observers, sensor fusion, nonlinear filters, strapdown systems, invariance, inertial navigation, extended Kalman filters.

Abstract: In this paper we propose a nonlinear observer (i.e. a "filter") for estimating the orientation of a flying rigid body, using measurements from low-cost inertial and magnetic sensors. It has by design a nice geometrical structure appealing from an engineering viewpoint; it is easy to tune, computationally very economic, and with guaranteed (at least local) convergence around every trajectory. Moreover it behaves sensibly in the presence of acceleration and magnetic disturbances. We illustrate its good performance on experimental comparisons with a commercial system, and demonstrate its simplicity by implementing it on a 8-bit microcontroller.

1 INTRODUCTION

Aircraft, especially Unmanned Aerial Vehicles (UAV), commonly 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 –often called Attitude Heading Reference Systems (AHRS)– rely on light and cheap "strapdown" gyroscopes, accelerometers and magnetometers. The various measurements are "merged" according to the motion equations of the aircraft assuming a flat non-rotating Earth, usually with some kind of Extended Kalman Filter (EKF). For more details about avionics, various inertial navigation systems and sensor fusion, see for instance (Collinson, 2003; Kayton and Fried, 1997; Grewal et al., 2007) and the references therein.

While the EKF is a general method capable of good performance when properly tuned, it suffers several drawbacks: it is not easy to choose the numerous parameters; it is computationally expensive, which is a problem in low-cost embedded systems; it is usually difficult to prove the convergence, even at first-order, and the designer has to rely on extensive simulations.

An alternative route is to use an ad hoc nonlinear observer as proposed in (Thienel and Sanner, 2003; Mahony et al., 2005; Hamel and Mahony, 2006; Martin and Salaün, 2007; Mahony et al., 2008). In the absence of a general method the main difficulty is of course to find such an observer. In this paper we use the rich geometric structure of the attitude-heading problem to derive an observer by the method developed in (Bonnabel et al., 2007), building up on the preliminary work (Martin and Salaün, 2007). It has by design a nice geometrical structure appealing from an engineering viewpoint; it is easy to tune, computationally very economic, and with guaranteed (at least local) convergence around every trajectory. Moreover it behaves sensibly in the presence of acceleration and magnetic disturbances. We illustrate its good performance on experimental comparisons with the commercial Microbotics MIDG II system, and demonstrate its simplicity by implementing it on a 8-bit Atmel AVR microcontroller.

2 THE PHYSICAL SYSTEM

2.1 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 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
- $A = ge_3$ is the (constant) gravity vector in North-East-Down coordinates (the unit vectors e_1, e_2, e_3 point respectively North, East, Down)
- a is the specific acceleration vector, in this case the aerodynamic 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 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 see (Stevens and Lewis, 2003) or any other good textbook on aircraft modeling, and section 7 for useful formulas used in this paper.

2.2 Measurements

We use three triaxial sensors providing nine scalar measurements: 3 gyros measure $\omega_m = \omega + \omega_b$, where ω_b is a constant vector bias; 3 magnetometers measure $y_B = b_s q^{-1} * B * q$, where $B = B_1 e_1 + B_3 e_3$ is the Earth magnetic field in NED coordinates and $b_s > 0$ is a constant scaling factor; 3 accelerometers measure $a_m = a_s a$, where $a_s > 0$ is a constant scaling factor. As customary in low-cost unaided attitude heading systems we assume the linear acceleration $\dot{V}$ small, hence approximate the accelerometers measurements by $y_A := a_s q^{-1} * A * q$ (the sign is reversed for convenience). All the nine measurements are of course also corrupted by noise.

There is some freedom in modeling the sensors imperfections. A simple first-order observability analysis reveals that up to six unknown constants can be estimated: besides the gyro bias ω_b, it is possible to estimate two imperfections on y_B and one on a_m, or one on y_B and two on a_m. Nevertheless it is impossible to model three imperfections on a_m: in particular if we write $a_m = a + a_b$, with a_b a constant vector bias, only two components of a_b are observable; moreover, only one imperfection on a_m can be estimated without relying on the possibly disturbed magnetic measurements. On the other hand it is also impossible to estimate the three components of the magnetic field B, but only the North and Down components.

In an AHRS it is usually desirable to use the magnetic measurements to estimate only the heading, so that a magnetic disturbance does not affect the estimated attitude. We will see this can be achieved by considering $y_C := y_A \times y_B = c_s q^{-1} * C * q$, where $c_s := a_s b_s > 0$ and $C := A \times B$, rather than the direct measurement y_B. Notice that $\langle y_A, y_C \rangle = \langle A, C \rangle = 0$, so that we are left with 8 independent measurements; as a consequence only five unknown constants can now be estimated. This is not a drawback and is even beneficial since the observer will then not depend on the latitude-varying B_3.

2.3 The Model

To design our observer we thus consider the system

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

$$\dot{\omega}_b = 0 \tag{2}$$

$$\dot{a}_s = 0 \tag{3}$$

$$\dot{c}_s = 0 \tag{4}$$

with the output

$$\begin{pmatrix} y_A \\ y_C \end{pmatrix} = \begin{pmatrix} a_s q^{-1} * A * q \\ c_s q^{-1} * C * q \end{pmatrix}. \tag{5}$$

This system is observable since all the state variables can be recovered from the known quantities ω_m, y_A, y_C and their derivatives: from (5), $a_s = \frac{1}{g} \|y_A\|$ and $c_s = \frac{1}{B_1 g} \|y_C\|$; hence we know the action of q on the two independent vectors A and C, which completely defines q in function of y_A, y_C, a_s, c_s. Finally (1) yields $\omega_b = \omega_m - 2q^{-1}\dot{q}$.

3 THE NONLINEAR OBSERVER

3.1 Invariance of the System Equations

There is no general method for designing a nonlinear observer for a given system. When the system has a rich geometric structure the recent paper (Bonnabel et al., 2007) provides a constructive method to this problem. In short, when a system of state dimension n is invariant by a transformation group of dimension $r \le n$, the method yields the general invariant preobserver. More importantly the error between the estimated and actual states is described in suitable invariant coordinates by a system of dimension $2n - r$; in particular the error system has dimension n when the group has dimension n. This result, which is reminiscent of the linear case, greatly simplifies the convergence analysis.

In our case the transformation generated by constant rotations and translations in the body-fixed frame and output scaling

$$\varphi_{(q_0,\omega_0,a_0,c_0)}\begin{pmatrix} q\\ \omega_b\\ a_s\\ c_s\end{pmatrix} = \begin{pmatrix} q*q_0\\ q_0^{-1}*\omega_b*q_0+\omega_0\\ a_0a_s\\ c_0c_s\end{pmatrix}$$

$$\psi_{(q_0,\omega_0,a_0,c_0)}(\omega_m) = q_0^{-1}*\omega_m*q_0+\omega_0$$

$$\rho_{(q_0,\omega_0,a_0,c_0)}\begin{pmatrix} y_A\\ y_C\end{pmatrix} = \begin{pmatrix} a_0q_0^{-1}*y_A*q_0\\ c_0q_0^{-1}*y_C*q_0\end{pmatrix},$$

where q_0 is a unit quaternion, ω_0 a vector in $\mathbb{R}^3$ and $a_0, c_0 > 0$, is easily seen to be a transformation group with the same dimension as the system (1)–(4).

The system (1)–(4) is indeed invariant by the transformation group since

$$\dot{\overbrace{q*q_0}} = \dot{q}*q_0 = \frac{1}{2}(q*q_0)*\big((q_0^{-1}*\omega_m*q_0+\omega_0) - (q_0^{-1}*\omega_b*q_0+\omega_0)\big)$$

$$\dot{\overbrace{q_0^{-1}*\omega_b*q_0+\omega_0}} = q_0^{-1}*\dot{\omega}_b*q_0 = 0$$

$$\dot{\overbrace{a_0a_s}} = a_0\dot{a}_s = 0$$

$$\dot{\overbrace{c_0c_s}} = c_0\dot{c}_s = 0.$$

Notice also that from a physical and engineering viewpoint, it is perfectly sensible for an observer using measurements expressed in the body-fixed frame not to be affected by the actual choice of 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 and output scalings should not affect the observer. This is precisely what the method of (Bonnabel et al., 2007) achieves.

3.2 The General Invariant Observer

Following the theory developed in (Bonnabel et al., 2007), see also (Martin and Salaün, 2007) for details, the general invariant observer writes

$$\dot{\hat{q}} = \frac{1}{2}\hat{q}*(\omega_m-\hat{\omega}_b)+(LE)*\hat{q} \tag{6}$$

$$\dot{\hat{\omega}}_b = \hat{q}^{-1}*(ME)*\hat{q} \tag{7}$$

$$\dot{\hat{a}}_s = \hat{a}_sNE \tag{8}$$

$$\dot{\hat{c}}_s = \hat{c}_sOE. \tag{9}$$

Here the invariant output error E is the 5×1 vector

$$E := \Big(\langle E_A,e_1\rangle,\langle E_A,e_2\rangle,\langle E_A,e_3\rangle,\langle E_C,e_1\rangle,\langle E_C,e_2\rangle\Big)^T$$

made up of the projections of the vectors

$$E_A := A-\frac{1}{\hat{a}_s}\hat{q}*y_A*\hat{q}^{-1}$$

$$E_C := C-\frac{1}{\hat{c}_s}\hat{q}*y_C*\hat{q}^{-1}.$$

Only 5 of the 6 possible projections are independent since $\langle A,C\rangle = 0$ and $\langle y_A,y_C\rangle = 0$ imply

$$\langle E_A,E_C\rangle = \langle A,E_C\rangle+\langle E_A,C\rangle;$$

L,M are 3×5 matrices and N,O are 1×5 matrices with entries possibly depending on the components of E and of the complete invariant I defined by

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

It is easy to check this observer is invariant. Notice also the two built-in desirable geometric features: $\|\hat{q}(t)\| = \|\hat{q}(0)\| = 1$ since LE is a vector of $\mathbb{R}^3$ (see section 7); $\hat{a}_s(t),\hat{c}_s(t) > 0$ provided $\hat{a}_s(0),\hat{c}_s(0) > 0$.

3.3 The Invariant Error System

Following (Bonnabel et al., 2007), an invariant state error is given by

$$\begin{pmatrix}\eta\\ \beta\\ \alpha\\ \gamma\end{pmatrix} = \begin{pmatrix}\hat{q}*q^{-1}\\ q*(\hat{\omega}_b-\omega_b)*q^{-1}\\ \frac{a_s}{\hat{a}_s}\\ \frac{c_s}{\hat{c}_s}\end{pmatrix}.$$

Therefore,

$$\dot{\eta} = \dot{\hat{q}}*q^{-1}-\hat{q}*(q^{-1}*\dot{q}*q^{-1}) = (LE)*\eta-\frac{1}{2}\eta*\beta$$

$$\begin{aligned}\dot{\beta} &= q*(\dot{\hat{\omega}}_b-\dot{\omega}_b)*q^{-1}+\dot{q}*(\hat{\omega}_b-\omega_b)*q^{-1}\\ &\quad -q*(\hat{\omega}_b-\omega_b)*q^{-1}*\dot{q}*q^{-1}\\ &= (\eta^{-1}*I*\eta)\times\beta+\eta^{-1}*(ME)*\eta\end{aligned}$$

$$\dot{\alpha} = -\frac{a_s\dot{\hat{a}}_s}{\hat{a}_s^2} = -\alpha NE$$

$$\dot{\gamma} = -\frac{c_s\dot{\hat{c}}_s}{\hat{c}_s^2} = -\gamma OE.$$

Since E is obtained from

$$E_A = A-\frac{a_s}{\hat{a}_s}\hat{q}*(q^{-1}*A*q)*\hat{q}^{-1} = A-\alpha\eta*A*\eta^{-1}$$

$$E_C = C-\gamma\eta*C*\eta^{-1}$$

we find as expected that the error system

$$\dot{\eta} = (LE)*\eta-\frac{1}{2}\eta*\beta \tag{10}$$

$$\dot{\beta} = (\eta^{-1}*I*\eta)\times\beta+\eta^{-1}*(ME)*\eta \tag{11}$$

$$\dot{\alpha} = -\alpha NE \tag{12}$$

$$\dot{\gamma} = -\gamma OE \tag{13}$$

depends only on the invariant state error $(\eta,\beta,\alpha,\gamma)$ and the "free" known invariant I, but not on the trajectory of the observed system (1)–(4). This property greatly simplifies the convergence analysis of the observer.

The linearized error system around the no-error equilibrium point $(\overline{\eta},\overline{\beta},\overline{\alpha},\overline{\gamma}) = (1,0,1,1)$ then reads

$$\delta\dot{\eta} = L\delta E - \frac{1}{2}\delta\beta \tag{14}$$

$$\delta\dot{\beta} = I\times\delta\beta + M\delta E \tag{15}$$

$$\delta\dot{\alpha} = -N\delta E \tag{16}$$

$$\delta\dot{\gamma} = -O\delta E, \tag{17}$$

where δE is the 5×1 vector

$$\Big(\langle\delta E_A, e_1\rangle, \langle\delta E_A, e_2\rangle, \langle\delta E_A, e_3\rangle, \langle\delta E_C, e_1\rangle, \langle\delta E_C, e_2\rangle\Big)^T = g\big(-2\delta\eta_2, 2\delta\eta_1, -\delta\alpha, 2B_1\delta\eta_3, -B_1\delta\gamma\big)^T$$

made up from the projections of the vectors

$$\delta E_A = A*\delta\eta - \delta\eta * A - \delta\alpha A = 2A\times\delta\eta - \delta\alpha A$$
$$\delta E_C = 2C\times\delta\eta - \delta\gamma C.$$

4 DESIGN OF L, M, N, O

Up to now, we have only investigated the structure of the observer. We now must choose the gain matrices L,M,N,O to meet the following requirements:

- the error must converge to zero, at least locally
- the local error behavior should be easily tunable, if possible with a clear physical interpretation
- the magnetic measurements should not affect the attitude estimate, but only the heading
- the behavior in the face of acceleration and/or magnetic disturbances should be sensible and understandable.

4.1 Local Design

It turns out that the previous requirements can easily be met *locally around every trajectory* by taking

$$L := \frac{1}{2g}\begin{pmatrix} 0 & -l_1 & 0 & 0 & 0\\ l_2 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -\frac{1}{B_1}l_3 & 0\end{pmatrix}$$

$$M := \frac{1}{2g}\begin{pmatrix} 0 & m_1 & 0 & 0 & 0\\ -m_2 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1}{B_1}m_3 & 0\end{pmatrix}$$

$$N := \frac{1}{g}\begin{pmatrix} 0 & 0 & -n & 0 & 0\end{pmatrix}$$

$$O := \frac{1}{B_1 g}\begin{pmatrix} 0 & 0 & 0 & 0 & -o\end{pmatrix}$$

for all constant $l_1,l_2,l_3,m_1,m_2,m_3,n,o > 0$. This will follow from the very simple form of the linearized error system (14)–(17). We insist that it is not usually obvious to come up with a similar convergence result for an EKF.

Indeed, (14)–(17) now reads

$$\delta\dot{\eta} = \mathrm{D_l}\delta\eta - \frac{1}{2}\delta\beta \tag{18}$$

$$\delta\dot{\beta} = \mathrm{D_m}\delta\eta + I\times\delta\beta \tag{19}$$

$$\delta\dot{\alpha} = -n\delta\alpha \tag{20}$$

$$\delta\dot{\gamma} = -o\delta\gamma \tag{21}$$

where

$$\mathrm{D_l} = \begin{pmatrix} -l_1 & 0 & 0\\ 0 & -l_2 & 0\\ 0 & 0 & -l_3\end{pmatrix}$$

$$\mathrm{D_m} = \begin{pmatrix} m_1 & 0 & 0\\ 0 & m_2 & 0\\ 0 & 0 & m_3\end{pmatrix}.$$

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

- the longitudinal subsystem
$$\begin{pmatrix}\delta\dot{\eta}_2\\ \delta\dot{\beta}_1\end{pmatrix} = \begin{pmatrix} -l_1 & -\frac{1}{2}\\ 0 & m_1\end{pmatrix}\begin{pmatrix}\delta\eta_2\\ \delta\beta_1\end{pmatrix}$$
- the lateral subsystem
$$\begin{pmatrix}\delta\dot{\eta}_1\\ \delta\dot{\beta}_2\end{pmatrix} = \begin{pmatrix} -l_2 & -\frac{1}{2}\\ 0 & m_2\end{pmatrix}\begin{pmatrix}\delta\eta_1\\ \delta\beta_2\end{pmatrix}$$
- the heading subsystem
$$\begin{pmatrix}\delta\dot{\eta}_3\\ \delta\dot{\beta}_3\end{pmatrix} = \begin{pmatrix} -l_3 & -\frac{1}{2}\\ 0 & m_3\end{pmatrix}\begin{pmatrix}\delta\eta_3\\ \delta\beta_3\end{pmatrix}$$
- the scaling subsystem
$$\delta\dot{\alpha} = -n\delta\alpha$$
$$\delta\dot{\gamma} = -o\delta\gamma.$$

When $I \neq 0$ the longitudinal, lateral and heading subsystems are slightly coupled by the biases errors $\delta\beta$.

We now prove $(\delta\eta,\delta\beta,\delta\alpha,\delta\gamma)\to(0,0,0,0)$ whatever $(l_1,l_2,l_3,m_1,m_2,m_3,n,o) > 0$. The scaling subsystem obviously converges. For the other variables we consider the Lyapunov function

$$V = \frac{l_1}{2}\delta\eta_1^2 + \frac{l_2}{2}\delta\eta_2^2 + \frac{l_3}{2}\delta\eta_3^2 + \frac{1}{4}\|\delta\beta\|^2.$$

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

$$\dot{V} = -\left(l_1 m_1\delta\eta_1^2 + l_2 m_2\delta\eta_2^2 + l_3 m_3\delta\eta_3^2\right) \leq 0.$$

Since V is bounded from below, this 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

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

Therefore $\delta\eta(t)$ and $\delta\beta(t)$ are bounded. Equation (18) implies that $\delta\dot\eta(t)$ is bounded too, and finally that $\ddot V$ is bounded. Hence $\dot V$ is uniformly continuous and by Barbalat's lemma

$$\dot V \to 0 \;\Rightarrow\; \delta\eta \to 0.$$

Integrating (18), we get

$$\begin{aligned}\int_0^t \delta\dot\eta(\tau)d\tau &= \delta\eta(t)-\delta\eta(0)\\ &= \int_0^t (D_l\delta\eta(\tau)-\frac{1}{2}\delta\beta(\tau))d\tau.\end{aligned}$$

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

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

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

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

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

4.2 Global Design

We now look for correction terms ensuring global convergence whereas preserving the previous nice first-order properties. It is useful to define the following vectors and error output:

$$\begin{aligned}D &= C\times A\\ y_D &= y_C\times y_A\\ E_D &= D-\frac{1}{\hat a_s\hat c_s}\hat q * y_D * \hat q^{-1}.\end{aligned}$$

If we define the matrix L by

$$LE := \frac{l_a}{g^2}A\times E_A + \frac{l_c}{(B_1 g)^2}C\times E_C + \frac{l_d}{(B_1 g^2)^2}D\times E_D$$

it is easy to see that the zero-order part of L is the same as the constant L of section 4.1 with

$$\begin{aligned}l_1 &= l_a + l_c\\ l_2 &= l_a + l_d\\ l_3 &= l_c + l_d.\end{aligned}$$

To find a candidate Lyapunov function we also select the matrices (M,N,O) by

$$\begin{aligned}ME &:= \sigma LE\\ NE &:= n\big(\frac{l_a}{g^2}E_A^T(E_A-A)+\frac{l_d}{(B_1g^2)^2}E_D^T(E_D-D)\big)\\ OE &:= o\big(\frac{l_c}{(B_1g)^2}E_C^T(E_C-C)+\frac{l_d}{(B_1g^2)^2}E_D^T(E_D-D)\big)\end{aligned}$$

with $(l_a,l_c,l_d,\sigma,n,o)>0$. The positive function $V := \frac{1}{\sigma}\|\beta\|^2+\frac{l_a}{g^2}\|E_A\|^2+\frac{l_c}{(B_1g)^2}\|E_C\|^2+\frac{l_d}{(B_1g^2)^2}\|E_D\|^2$ satisfies $\dot V\leq 0$. The convergence proof follows the main lines of (Hamel and Mahony, 2006), with added technicalities.

With these choices the first-order behavior is the same as in section 4.1. The difference is the existence of the proportional factor σ between the $l_i's$, $m_i's$ coefficients. We now have only 4 independent gains for the lateral, longitudinal and heading subsystems instead of 6. We choose to fix first the natural frequencies of the lateral (ω_l), longitudinal (ω_L) and heading (ω_h) subsystems. There is only 1 parameter left to fix the free damping ratios.

5 EFFECTS OF DISTURBANCES

Two main disturbances may affect the model. When $\dot V \neq 0$, the accelerometers measure in fact $a_s q^{-1} * A^* * q$ where $A^* := -\dot V + A$. Magnetic disturbances will also change B into some B^*. For simplicity we consider that A^*, B^* are constant. The measured outputs now become

$$\begin{pmatrix} y_{A^*}\\ y_{C^*}\end{pmatrix} = \begin{pmatrix} a_s q^{-1} * A^* * q\\ c_s q^{-1} * C^* * q\end{pmatrix}.$$

The error system is unchanged but E is now the 5×1 vector

$$E := \Big(\langle E_A,e_1\rangle,\langle E_A,e_2\rangle,\langle E_A,e_3\rangle,\langle E_C,e_1\rangle,\langle E_C,e_2\rangle\Big)^T$$

made up of the projections of the vectors

$$\begin{aligned}E_A &:= A-\frac{1}{\hat a_s}\hat q * y_{A^*} * \hat q^{-1}\\ E_C &:= C-\frac{1}{\hat c_s}\hat q * y_{C^*} * \hat q^{-1}.\end{aligned}$$

Let us define the points $(\overline\eta,\overline\beta,\overline\alpha,\overline\gamma)$ as following

$$\begin{aligned}\overline\beta &= 0\\ \overline\eta * A^* * \overline\eta^{-1} &= (0\ 0\ \|A^*\|)\\ \overline\eta * C^* * \overline\eta^{-1} &= (0\ \|C^*\|\ 0)\\ \overline\alpha = \frac{\|A^*\|}{\|A\|} &\quad\text{and}\quad \overline\gamma = \frac{\|C^*\|}{\|C\|}\end{aligned}$$

Doing the frame rotation defined by $\overline{\eta}$ we can define the new variables

$$\tilde{\eta} = \eta * \overline{\eta}^{-1} \qquad \tilde{\beta} = \overline{\eta} * \beta * \overline{\eta}^{-1}$$
$$\tilde{\alpha} = \alpha\overline{\alpha} \qquad \tilde{\gamma} = \gamma\overline{\gamma}.$$

The error system with these new variables writes

$$\dot{\tilde{\eta}} = -\frac{1}{2}\tilde{\eta} * \tilde{\beta} + (\tilde{L}\tilde{E}) * \tilde{\eta}$$
$$\dot{\tilde{\beta}} = (\tilde{\eta}^{-1} * \tilde{I} * \tilde{\eta}) \times \tilde{\beta} + \tilde{\eta}^{-1} * (\tilde{M}\tilde{E}) * \tilde{\eta}$$
$$\dot{\tilde{\alpha}} = -\tilde{\alpha}\tilde{N}\tilde{E}$$
$$\dot{\tilde{\gamma}} = -\tilde{\gamma}\tilde{O}\tilde{E}$$

where the new output error $\tilde{E}$ is made up of the projections of the vectors

$$\tilde{E}_A = A - \tilde{\alpha}\tilde{\eta} * A * \tilde{\eta}^{-1}$$
$$\tilde{E}_C = C - \tilde{\alpha}\tilde{\eta} * C * \tilde{\eta}^{-1}.$$

So $(\tilde{\eta}, \tilde{\beta}, \tilde{\alpha}, \tilde{\gamma})$ verify the same error system as $(\eta, \beta, \alpha, \gamma)$. In the new frame (A^*, C^*) play the same role as (A, C). All the properties of the observer are therefore preserved.

An important case is when only the magnetic field is perturbed, where we consider A and $C^* = (C_1^* \; C_2^* \; 0)$ (instead of $C = (0 \; gB_1 \; 0)$). Expliciting the new equilibrium point $(\overline{\eta}, \overline{\beta}, \overline{\alpha}, \overline{\gamma})$ of the error system it can be seen that

$$\overline{\phi} = \overline{\theta} = 0 \text{ and } \overline{\psi} = \arctan\frac{C_1^*}{C_2^*}$$
$$\overline{\beta} = \overline{\alpha} = 0 \text{ and } \overline{\gamma} = \frac{\|C^*\|}{\|C\|},$$

where $(\overline{\phi}, \overline{\theta}, \overline{\psi})$ are the Euler angles corresponding to $\overline{\eta}$. In particular only the yaw angle $\overline{\psi}$ and $\overline{\gamma}$ are affected by the magnetic disturbance.

6 EXPERIMENTAL VALIDATION

We now compare the behavior of our observer with the commercial Microbotics MIDG II system used in Vertical Gyro mode. The following results have been obtained with the observer

$$\dot{\hat{q}} = \frac{1}{2}\hat{q} * (\omega_m - \hat{\omega}_b) + (LE) * \hat{q} + k(1 - \|\hat{q}\|^2)\hat{q}$$
$$\dot{\hat{\omega}}_b = \hat{q}^{-1} * (ME) * \hat{q}$$
$$\dot{\hat{a}}_s = \hat{a}_s NE$$
$$\dot{\hat{c}}_s = \hat{c}_s OE$$

and the choice of matrices defined by the parameters below. 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.

We feed the observer with the raw measurements from the MIDG II gyros, acceleros and magnetic sensors. The observer is implemented in Matlab Simulink and its values are compared to the MIDG II results (computed according to the user manual by some Kalman filter). In order to have similar behaviors, we have chosen

$$l_a = 6e-2 \qquad l_c = 1e-1 \qquad l_d = 6e-2$$
$$m_a = 3.2e-3 \qquad m_c = 5.3e-3 \qquad m_d = 3.2e-3$$
$$n = 0.25 \qquad o = 0.5.$$

6.1 Comparison with a Commercial Device (Figure 1)

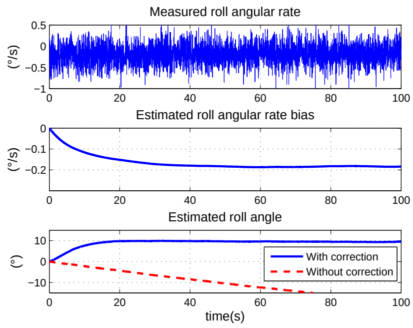
We first want to put in evidence the different properties of our observer we mentioned before. Therefore we did a long experience which can be divided into 3 parts:

- for $t < 240s$ the system is left at rest until the biases reach constant values. Fig.1(a) highlights the importance of the correction term in the angle estimation: without correction the estimated roll angle diverges with a slope of $-0.18°/s$ (bottom plot), which is indeed the final value of the estimated bias (middle plot) (Fig.1(a) and 1(b)).
- for $240s < t < 293s$ we move the system in all directions. The observer and the MDG II give very similar results (Fig.1(c)).
- at $t = 385s$ the system is motionless and a magnet is put close to the sensors for 10s. As expected only the estimated yaw angle is affected by the magnetic disturbance (Fig.1(d)); the MIDG II exhibits a similar behavior.

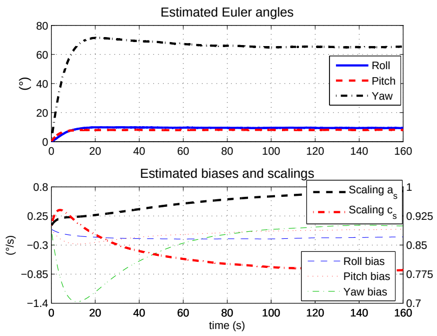
6.2 Influence of the Observer Correction Terms (Figure 2)

We have chosen the correction terms so as the magnetic measurements correct essentially the yaw angle and its corresponding bias, whereas the accelerometers measurements act on the other variables. We highlight this property on the following experiment (Fig.2). Once the biases have reached constant values, the system is left at rest during 35 minutes:

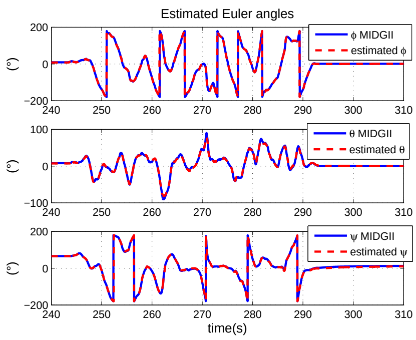
- for $t < 600s$ the results are very similar for the observer and the MIDG II.
- at $t = 600s$ the "magnetic correction terms" are switched off, i.e. the gains l_c,l_d,m_c,m_d and o are



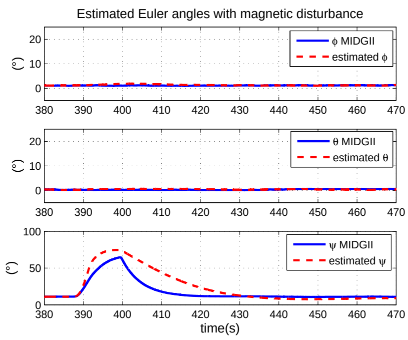
(a) Motionless estimated roll angle.



(b) Motionless estimated gyros biases and Euler angles.



(c) Dynamic estimated Euler angles.



(d) Estimated Euler angles with magnetic disturbances.

Figure 1: Experimental validation using Matlab.

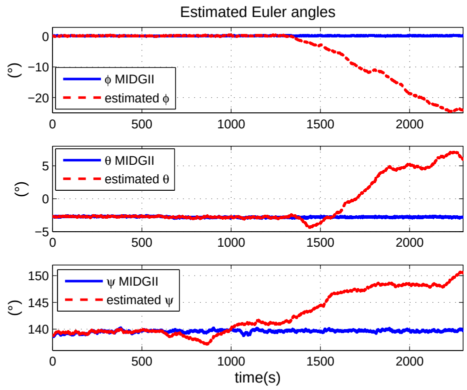


Figure 2: Influence of correction terms.

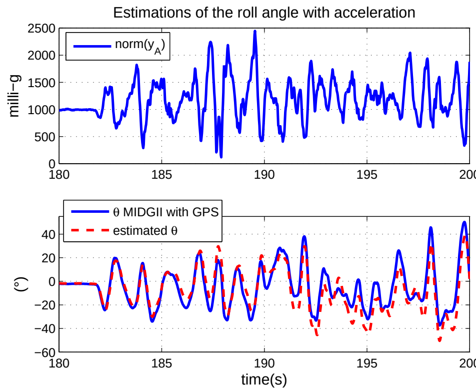


Figure 3: Estimated roll angle at $\dot{V} \neq 0$.

set to 0. The yaw angle estimated by the observer diverges because the corresponding bias is not perfectly estimated. Indeed, these variables are not observable without the magnetic measurements. The other variables are not affected.

- at $t = 1300s$ the "accelerometers correction terms" are also switched off, i.e. l_a,m_a and n are set to 0. All the estimated angles now diverge.

6.3 Acceleration Disturbance: $\dot{V} \neq 0$ (Figure 3)

The hypothesis $\dot{V} = 0$ may be wrong. In this case the observer does not converge any more to the right values. Indeed we illustrate this point on the figure 3 by comparing the roll angle estimated by our observer and the roll angle estimated by the MIDG II in INS mode (in this mode the attitude and heading estimations are aided by a GPS engine, hence are close to the "true" values).

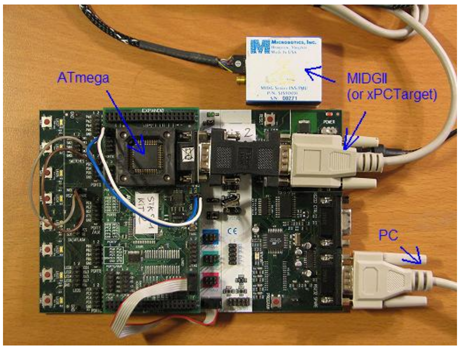


Figure 4: Experimental protocol.

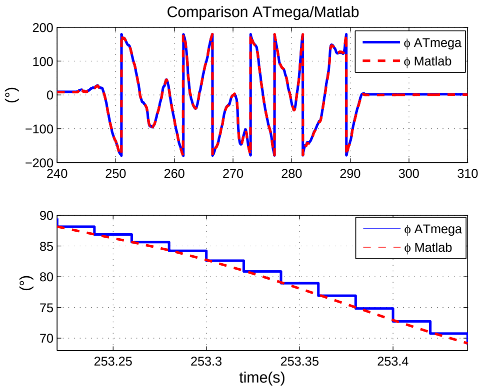


Figure 5: Comparison ATmega/Matlab.

7 IMPLEMENTATION ON A 8-BIT MICROCONTROLLER

In the preceding section we validated the Matlab Simulink implementation of our observer. To demonstrate its computational simplicity we have also implemented it on a 8-bit microcontroller (Atmel ATmega128 running at 11.0592MHz on development kit STK500/501). The computations are done in C with the standard floating point emulation. The microcontroller is fed with the MIDG II raw data at a 50Hz rate. We have used a simple Euler explicit approximation for the integration scheme.

The experimental protocol was the following:

1. move the sensors in all directions and save the MIDG II raw measurements and estimations at 50Hz.
2. use of xPCTarget to feed the ATmega with these data at 50Hz via a serial port and send at the same time the estimated variables. given by the microcontroller via a serial port to a computer
3. save these estimations.
4. compare offline the estimations given by the microcontroller and those given by the observer written in the Matlab embedded function.

This protocol can be illustrated by the figure 4 where xPCTarget has been replaced by the MIDG II.

We obtain the results of the figure 5. The two estimations are very similar. We see on the bottom plot the discretization at 50Hz due to the microcontroller.

REFERENCES

Bonnabel, S., Martin, P., and Rouchon, P. (2007). Invariant observers. *arxiv.math.OC/0612193*. Accepted for publication in IEEE Trans. Automat. Control.

Collinson, R. (2003). *Introduction to avionics systems*. Kluwer Academic Publishers, second edition.

Grewal, M., Weill, L., and Andrews, A. (2007). *Global positioning systems, inertial navigation, and integration*. Wiley, second edition.

Hamel, T. and Mahony, R. (2006). Attitude estimation on $SO(3)$ based on direct inertial measurements. In *Proc. of the 2006 IEEE International Conference on Robotics and Automation*, pages 2170–2175.

Kayton, M. and Fried, W., editors (1997). *Avionics navigation systems*. Wiley, second edition.

Mahony, R., Hamel, T., and Pfimlin, J.-M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Trans. Automat. Control*. To appear.

Mahony, R., Hamel, T., and Pflimlin, J.-M. (2005). Complementary filter design on the special orthogonal group $SO(3)$. In *Proc. of the 44th IEEE Conf. on Decision and Control*, pages 1477–1484.

Martin, P. and Salaün, E. (2007). Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors. In *Proc. of the 46th IEEE Conf. on Decision and Control*.

Stevens, B. and Lewis, F. (2003). *Aircraft control and simulation*. Wiley, second edition.

Thienel, J. and Sanner, R. (2003). 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.

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 $\in \mathbb{R}^3$. 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.