Three examples of the stability properties of the invariant extended Kalman filter
Axel Barrau, Silvere Bonnabel

To cite this version:
Axel Barrau, Silvere Bonnabel. Three examples of the stability properties of the invariant extended Kalman filter. 2017. hal-01526467

HAL Id: hal-01526467
https://inria.hal.science/hal-01526467
Preprint submitted on 23 May 2017

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.
Three examples of the stability properties of the invariant extended Kalman filter

A. Barrau* S. Bonnabel**

* SAFRAN TECH, Groupe Safran, Rue des Jeunes Bois - Châteaufort, 78772 Magny Les Hameaux CEDEX.
** MINES ParisTech, PSL Research University, Centre for Robotics, 60 bd Saint-Michel, 75006 Paris, France.

Abstract: In the aerospace industry the (multiplicative) extended Kalman filter (EKF) is the most common method for sensor fusion for guidance and navigation. However, from a theoretical point of view, the EKF has been shown to possess local convergence properties only under restrictive assumptions. In a recent paper, we proved a slight variant of the EKF, namely the invariant extended Kalman filter (IEKF), when used as a nonlinear observer, possesses local convergence properties under the same assumptions as those of the linear case, for a class of systems defined on Lie groups. This is especially interesting as the IEKF also retains all the desirable features of the standard EKF, especially its relevant tuning in the presence of noises. In the present paper we provide three examples of engineering interest where the theory is shown to apply, yielding three EKF-like algorithms with guaranteed local convergence properties. Beyond those contributions, the present article is sufficiently accessible to help the practitioner understand through concrete examples the general IEKF theory, and to provide him with guidelines for the design of IEKFs.

Keywords: Estimation, Kalman filtering, nonlinear systems, Lie groups, navigation.

1. INTRODUCTION

In the aerospace industry the (multiplicative) extended Kalman filter (EKF) is the most popular method for sensor fusion for guidance and navigation. However, from a theoretical point of view, the EKF has been shown to possess local convergence properties only under restrictive assumptions Song and Grizzle (1995); Krener (2003), and as a matter of fact it can actually diverge, even for small initial errors. The recent paper Barrau and Bonnabel (2017) proves a slight variant of the EKF, namely the invariant extended Kalman filter (IEKF), possesses local convergence properties under highly reasonable assumptions for a well characterized class of systems defined on Lie groups. The IEKF was originally introduced in Bonnabel (2007); Bonnabel et al. (2009b), and builds upon the theory of symmetry-preserving observers Bonnabel et al. (2009a). It can also be related to the generalized multiplicative EKF of Martin and Salaün (2010), the discrete EKF on Lie groups Bourmaud et al. (2013), see also de Ruiter and Forbes (2016).

The principles of the IEKF theory as presented in Barrau and Bonnabel (2017) are not easy to grasp. The main goal of the present paper is thus to provide a user friendly presentation and discussion of the IEKF as described in Barrau and Bonnabel (2017), and to illustrate its stability properties on three examples of engineering interest. Even though the purpose of the present article is essentially tutorial, it also contains novel theoretical results as we derive three non-linear filters for three examples of engineering interest, and guarantee stability of the derived filters.

Those examples could certainly be tackled through non-linear observers, along the lines of e.g., Hua et al. (2014); Wolfe et al. (2011); Batista et al. (2014); Izadi and Sanyal (2014); Sanyal and Nordkvist (2012); Zamani et al. (2014); Lee et al. (2007); Hua et al. (2016); Tayebi et al. (2007), and (almost) global convergence properties could be - or have already been - obtained. The interest (and difference) of our approach with respect to the non-linear observer literature though, is that the three non-linear proposed filters (namely IEKFs) 1- accomodate discrete time measurements with arbitrary and varying sampling times, 2- the gain tuning matches the modeled variance of the noises through (linearized) Kalman's theory 3- this implies the gains easily accomodate time-varying features, such as time-varying covariance matrices, 4- contrarily to non-linear observers, the filter provides an indication (through the covariance matrix $P_t$) of the extent of uncertainty conveyed by the estimate and 5- the filters viewed as observers, that is, when noise is turned off, converge around any trajectory, with an attraction radius which is uniform over time. It is worthy to note that, to that respect, the three filters achieve the same goals as the ones pursued by the very recent XKF Johansen and Fossen (2016), albeit a wholly different method. Note though, it has not yet been shown an XKF may be built on the following examples.

In a nutshell, the IEKFs proposed here should be appealing to the aerospace engineers: they retain all the characteristics of the standard EKF (first-order optimality, realtive ease of tuning, adaptivity to time-varying features...
and to discrete aperiodic measurements), but with additional
stability properties, when studied in a deterministic
setting using the tools of dynamical systems theory.

The paper is organized as follows. Section II is a concise
tutorial summary and discussion on Barrau and Bonnabel
(2017). Each following section deals with an example. A
conclusion seemed not necessary, so it was omitted.

2. REVIEW OF THE IEKF METHODOLOGY AND
CONVERGENCE PROPERTIES

In this section, we review the IEKF methodology as
presented in Barrau and Bonnabel (2017), that is, for
continuous time dynamics with discrete time observations
for systems defined on Lie groups. The exposure is meant
to be concise and tutorial, and is enhanced by discussions.
See the Appendix for more details on matrix Lie groups.

2.1 Considered class of systems and IEKF equations

Consider in this section a dynamics on a matrix Lie group
\( G \subset \mathbb{R}^{N \times N} \) with state \( \chi_t \in G \) satisfying:
\[
\frac{d}{dt} \chi_t = f_u(\chi_t) + \chi_t w_t,
\]
where \( w_t \) is a continuous white noise belonging to the Lie
algebra \( \mathfrak{g} \) (see the Appendix). Let \( q = \dim G \) denote the
dimension of the Lie group \( G \) (or alternatively defined by \( q = \dim \mathfrak{g} \)).
Assume moreover the following relation holds
\[
f_u(ab) = a f_u(b) + f_u(a) b - a f_u(b) a
\]
for all \((a, b) \in U \times G \times G\). This system can be associated
with two different kinds of discrete observations at
arbitrary times \( t_0 < t_1 < t_2, \ldots \).

Left-invariant observations

The first family of outputs we are interested in write:
\[
Y_{t_n}^1 = \chi_{t_n} \left(d^1 + B_n^1 \right) + V_{t_n}^1, \ldots, Y_{t_n}^k = \chi_{t_n} \left(d^k + B_n^k \right) + V_{t_n}^k,
\]
where \((d^1), \ldots, (d^k)\) are known vectors of \( \mathbb{R}^{N} \), and where the
\((V_{t_n}^1), \ldots, (V_{t_n}^k)\) are centered Gaussian variables noises
with known covariance matrices.

The outputs are said to be “left-invariant” as, in the
absence of noise, the outputs are of the form \( h(\chi) = \chi d \)
so that, \( \chi^2 h(\chi_1) = h(\chi_2 \chi_1) \). This property is also
referred to as left equivariance in the mathematics literature
and in the theory of symmetry-preserving observers.
For left-invariant observations, a Left-Invariant
EKF (LIEKF) should always be used.

The Left-Invariant Extended Kalman Filter (LIEKF) is
defined through the usual following propagation and update
steps:
\[
\frac{d}{dt} \hat{x}_t = f_u(\hat{x}_t), \quad t_{n-1} \leq t < t_n, \quad \text{Propagation} (4)
\]
\[
\hat{x}_{t_n}^+ = \hat{x}_{t_n} \exp \left[ L_n \begin{bmatrix} \hat{x}_{t_n}^{-1} Y_{t_n}^1 - d^1 \\ \vdots \\ \hat{x}_{t_n}^{-1} Y_{t_n}^k - d^k \end{bmatrix} \right], \quad \text{Update} \ (5)
\]
where the function \( L_n : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}^q \) is defined through
linearizations as in the conventional EKF theory. But here,
instead of considering the usual linear state error
\( \hat{x}_t - \chi_t \), one must consider the following left-invariant error
between true state \( \chi_t \) and the estimated state \( \hat{x}_t \):
\[
\eta_t^L = \chi_t^{-1} \hat{x}_t.
\]
which is the counterpart of the linear error \( \hat{x}_t - \chi_t \) (which
has no proper meaning in the present context), when dealing
with a state space that is a Lie group. Note that, this error is nominally equal to identity matrix and
not zero. The rationale of the IEKF theory, and more
generally the theory of symmetry-preserving observers, is
to linearize the error system at the propagation and update
state. It turns out the error system, with an error defined
this way, has remarkable properties, that are key to prove
the IEKF stability properties of Barrau and Bonnabel
(2017).

Right-invariant observations

The second family of observations we are interested in have the form
\[
Y_{t_n}^1 = \chi_{t_n}^{-1} (d^1 + V_{t_n}^1) + B_n^1, \ldots, Y_{t_n}^k = \chi_{t_n}^{-1} (d^k + V_{t_n}^k) + B_n^k,
\]
with the same notation as in the previous paragraph.

The Right-Invariant EKF (RIEKF), always be used for
right-invariant observations of the form (7) is defined here
in the same way, alternating between continuous time
propagation and discrete time update steps:
\[
\frac{d}{dt} \hat{x}_t = f_u(\hat{x}_t), \quad t_{n-1} \leq t < t_n, \quad \text{Propagation} (8)
\]
\[
\hat{x}_{t_n}^+ = \exp \left[ L_n \begin{bmatrix} \hat{x}_{t_n} Y_{t_n}^1 - d^1 \\ \vdots \\ \hat{x}_{t_n} Y_{t_n}^k - d^k \end{bmatrix} \right] \hat{x}_{t_n}, \quad \text{Update} \ (9)
\]
To tune the gain \( L_n \), the state error must be linearized, but
in this case we rather consider the right-invariant error
\[
\eta_t^R = \hat{x}_t \chi_t^{-1}.
\]

Gain tuning

To tune the gain matrix \( L_n \), one must linearize the error equation associated to (6), or respectively
(10). To do so, the user can refer to the general theory of
Barrau and Bonnabel (2017), or rather proceed to a case
derivation as done in the examples below, which is
recommended. In any case, one can associate to the non-
linear error (6), or (10), a vector \( \xi \in \mathbb{R}^q \) that captures the
error up to the first order. It can be used to obtain a linear
approximation to the true error system, of the form:
\[
\frac{d}{dt} \xi_t = A \xi_t + D(\hat{x}_t) \tilde{w}_t
\]
where \( \tilde{w}_t \) is a continuous noise in \( \mathbb{R}^q \) and to a linearized
error update equation of the form
\[
\xi_{t_n} = \xi_{t_n} - L_n (H \xi_{t_n} + E(\hat{x}_{t_n}) V_{t_n})
\]
with \( V_n \) a centered Gaussian. To account for the fact
the stochastic terms entering the system depend on the
estimated trajectory, we define as in the standard EKF
theory with non-additive noises (see e.g. Stengel (1986))
the covariance matrices
\[
Q(\hat{x}_t) = D(\hat{x}_t) \text{Cov}(\tilde{w}_t) D(\hat{x}_t)^T
\]
\[
N(\hat{x}_{t_n}) = E(\hat{x}_{t_n}) \text{Cov}(V_{t_n}) E(\hat{x}_{t_n})^T
\]
As in the standard EKF methodology, the “optimal” gain
\( L_n \) is then obtained through the Kalman equations:
\[
\frac{d}{dt} P_t = A_t P_t + P_t A_t^T + Q(\hat{x}_t),
\]
\[
S_n = H P_t H^T + N(\hat{x}_{t_n}),
\]
\[
L_n = P_{t_n} H^T S_n^{-1}, \quad P_{t_n} = (I - L_n H) P_{t_n}.
\]

2.2 Geometrical insight for the practitioner

In this section we provide the user with some (novel) geometrical insight about the EIKF linearization procedure, which may prove helpful, albeit not strictly necessary to apply the theory. Even if it is not apparent when looking at the original problem, the fact that the state space may be identified to a matrix Lie group \( G \), which is a subset of \( \mathbb{R}^{N \times N} \), means the state space is a curved space. Albeit not rigorous, one can think of the state space as a sphere. The error \( \eta_t \) of (6), resp. (10), is an element of \( G \), and is equal to \( I_N \) (only) when \( \hat{x}_t = x_t \). Thus, the error system is always linearized around the same point, that is, the identity matrix. To linearize it, the rationale (see Bonnabel et al. (2009a)) is to identify a small error \( \eta_t \) with an element of the tangent space at \( I_N \) (called the Lie algebra of \( G \), and denoted by \( \xi \)). In turn this element of the tangent space is identified to a matrix \( \xi_t \) (the Lie exponential map appearing in (5), resp. (9)), in turn a way to map this tangent vector at \( I_N \) (the correction term) back to the group \( G \), and pre-multiplication in (5), resp. post-multiplication in (9), by \( \hat{x}_{t_n} \), is a way to apply this correction to the latest estimate \( \hat{x}_{t_n} \in G \). This is illustrated on the following cartoon.

2.3 IEKF general stability properties

Consider deterministic non-linear continuous time systems with discrete time observations, that is, general systems of the form
\[
\frac{d}{dt} x_t = f(x_t, u_t), \quad y_{t_n} = h(x_{t_n}).
\]
Even under observability conditions, designing non-linear observers is always a challenge. Notably, local convergence around any trajectory is a very rare property to obtain (see e.g. Aghannan and Rouchon (2003) which deals with the apparently simple problem of estimating the velocity from position and acceleration measurements of a mechanical system). A simple way to design an observer that might yield good results, is to apply a conventional EKF to the system
\[
\frac{d}{dt} x_t = f(x_t, u_t), \quad y_{t_n} = h(x_{t_n}),
\]
considering the process noise and measurement noise covariance matrices as design parameters. But it turns out the obtained filter is by no means guaranteed to converge around any trajectory of the system, unless some rather strong assumptions hold, see e.g. Song and Grizzle (1995); Bonnabel and Slotine (2015). By contrast, for systems defined on matrix Lie groups of the form (1)-(3) or (1)-(7) with noise turned off, it turns out the EIKF used as a deterministic non-linear observer converges around any trajectory, as follows.

**Theorem 1.** Barrau and Bonnabel (2017) Consider the system (1) with noise turned off, i.e.,
\[
\frac{d}{dt} \chi_t = f_u(\chi_t)
\]
Either one can assume left-invariant discrete output measurements (3) with noise turned off, i.e.
\[
Y^1_{t_n} = \chi_{t_n} d_1, \ldots, \quad Y^k_{t_n} = \chi_{t_n} d_k,
\]
and use the Left-invariant EKF (4)-(5).

Or, one can assume right-invariant output measurements (7) with noise turned off, i.e.
\[
\hat{Y}^1_{t_n} = \chi_{t_n}^{-1} d_1, \ldots, \quad \hat{Y}^k_{t_n} = \chi_{t_n}^{-1} d_k,
\]
and use the Right-invariant EKF (8)-(9).

Let \( A_t, Q, N \) be defined as in the EIKF algorithm, and let \( \Phi_{t_0} \) denote the square matrix defined by \( \Phi_{t_0} = I_{p} \), \( \Phi_{t_0} = A_t \). For simplicity of notation let \( Q(\chi_t) := Q_t \), and \( N(\chi_{t_n}) := N_{t_n} \). Assume there exist \( \alpha_1, \alpha_2, \beta_1, \beta_2, \delta_1, \delta_2, \delta_3, M \) such that:

\[
(1) \quad (\Phi_{t_n}^{s+1})^T \Phi_{t_n}^{s+1} \geq \delta_1 I_p \geq 0,
\]
\[
(2) \quad \exists q \in \mathbb{N}^*, \forall s > 0, \exists G_s \in \mathbb{R}^{p \times q}, Q_s = G_s Q^T G_s^T \text{ where } Q_s \geq \delta_2 I_p \geq 0,
\]
\[
(3) \quad N_{t_n} \geq \delta_3 I_{N_{t_n}} \geq 0,
\]
\[
(4) \quad \alpha_1 I_p \leq \int_{t_{n-M}}^{t_n} (\Phi_{t_0}^{s})^T Q_s (\Phi_{t_0}^{s})^T \leq \alpha_2 I_p,
\]
\[
(5) \quad \beta_1 I_p \leq \sum_{i=n-M}^{n-1} (\Phi_{t_n}^{s+1})^T H_t N_{t_n-1} \leq \beta_2 I_p.
\]

Then, there exists \( \epsilon > 0 \) such that, for any \( t_0 \geq 0 \), if the distance between \( \chi_{t_0} \) and \( x_{t_0} \) is less than \( \epsilon \), then the distance to the estimate \( \hat{x}_t \) and the true state \( x_t \) tends to zero as \( t \to \infty \).

The theorem is proved in Barrau and Bonnabel (2017). The proof is technical, non-trivial, and long. It brings to bear the very rich structure of the class of systems considered. Contrarily to previous literature on EKF convergence, the main feature here is that no assumption is whatsoever made on the estimate \( \hat{x}_t \)’s behavior.

2.4 Discussion on the role of noises

The theorem holds for the noise-free system, whereas the matrices \( Q(\chi_t) \) and \( N(\chi_{t_n}) \) involved in the gain tuning (13) are based on the noise’s characteristics. This may seem a contradictory approach. But, proving the stochastic counterpart of Theorem 1 would definitely be out of reach, whereas the theory of deterministic dynamical systems allows using systematic and relatively basic tools for stability analysis. And proving that a non-linear filter, once tuned for an actually noisy system, enjoys deterministic convergence properties in the absence of noise, is certainly desirable and reassuring: a guarantee of local stability in
the absence of noise is a great indication that the filter should not diverge for sufficiently small errors and noises.

3. FIRST EXAMPLE: ATTITUDE ESTIMATION

In this section we consider the problem of attitude estimation with gyroimeters and observation of two vectors with known directions in the fixed frame. This problem has received a lot of attention over the past decades, and the literature is too broad to be covered here (see e.g. Mahony et al. (2008); Izadi and Sanyal (2014); Sanyal and Nordkvist (2012); Zamani et al. (2014); Lee et al. (2007)). We insist that this example is very simple, and that we included it essentially for tutorial reasons. However, note that, contrarily to the non-linear attitude estimation literature just mentioned, the following filter accommodates discrete-time observations, which can be relevant in practice (for instance a star tracking system in a satellite could use infrequent snapshots to save energy), and time varying model uncertainties (for instance the quasi-static assumption below for a UAV does not hold in dynamic flight phases, so that the noise variance on the gravity measurement should be temporally increased when maneuvers are made). Note also that we already published the equations of the IEKF for this problem, but we never proved its local stability properties.

**Model** Consider the attitude of a vehicle, represented by the rotation matrix \( R_t \in SO(3) \) mapping the coordinates of a vector expressed in the vehicle frame to its coordinates in the static frame. The vehicle is endowed with gyroimeters measuring an angular velocity \( \omega_t \). The dynamics read:

\[
\frac{d}{dt} R_t = R_t (\omega_t + w_t)_x
\]

where \((b)_x\) is the skew symmetric matrix associated to vector \( b \) and \( w_t \) is the gyroscopes’ noise vector. Note that, \((14)\) bears a resemblance to a linear system. The resemblance is yet wholly artificial as \( SO(3) \) is not a vector space. Consider as observations for \( t_0 < t_1 < t_2 < \cdots \)

\[
Y_n = (R_{t_0}^T g + V_n^g; R_{t_0}^T b + V_n^b),
\]

where \( g \) and \( b \) are two non-collinear vectors in \( \mathbb{R}^3 \) and \( V_n^g, V_n^b \) two centered noises in \( \mathbb{R}^3 \). Generally those two measurements are respectively associated to an accelerometer that is supposed to measure the earth gravity under the quasi-static hypothesis, and to a magnetometer that measures the earth magnetic field. Relation \((12)\) holds.

**Error equation** The outputs \((15)\) are right-invariant. Thus one should devise a right-invariant EKF. The RIEKF \((8)-(9)\) for the system \((14)-(15)\) is defined by:

\[
\frac{d}{dt} \hat{R}_t = \hat{R}_t (\omega_t)_x \quad \hat{R}_n^T = \exp(L_n [\hat{R}_t, Y_n - (g; b)]) \hat{R}_n,
\]

with the notation \( R(x_1; x_2) = (R x_1; R x_2) \) for \( R \in SO(3) \) aux \( x_1, x_2 \in \mathbb{R}^3 \). To compute the gains \( L_n \) we write the evolution of the right invariant error \((10)\): \( \eta_t = \hat{R}_t R_t^T \). As concerns the propagation step, using that \((\omega_t + w_t)_x = - (\omega_t)_x - (w_t)_x \), we have

\[
\frac{d}{dt} \eta = \left( \frac{d}{dt} \hat{R}_t \right) R_t^T + \hat{R}_t \left( \frac{d}{dt} R_t^T \right) = \hat{R}_t (\omega_t)_x R_t^T + \hat{R}_t (\omega_t + w_t)_x R_t^T = 0 - \hat{R}_t (w_t)_x R_t^T = - (\hat{R}_t w_t)_x \eta,
\]

since \( \hat{R}_t (x \times y) = (\hat{R}_t x \times \hat{R}_t y) \). The update step writes

\[
\eta_t^+ = \exp(L_n [\eta g + \hat{R}_t V_n^g; \eta b - b + \hat{R}_t V_n^b]) \eta_t
\]

Linearized error equation and gain tuning To linearize the error equation we introduce the linearized state error vector \( \xi_t \in \mathbb{R}^3 \) in the Lie algebra \( \mathfrak{so}(3) \), defined by \( \eta_t^+ = \exp(\xi_t) = I_3 + \mathcal{L}_{\mathfrak{so}(3)}(\xi_t) + O(\|\xi_t\|^2) = I_3 + (\xi_t)_x + O(\|\xi_t\|^2) \). See the appendix for more details. Replacing \( \eta_t \) with \( I_3 + (\xi_t)_x \), the error propagation equation above reads

\[
\frac{d}{dt} (\xi_t)_x = - (\hat{R}_t w_t)_x (I_3 + (\xi_t)_x) = -(\hat{R}_t w_t)_x
\]

where we neglected terms of magnitude \( \|w_t\|\|\xi_t\| \), as in the standard EKF methodology in the presence of non-additive noises Stengel (1986) p 386. This yields the linearized equation

\[
\frac{d}{dt} \xi_t = 0 - \hat{R}_t w_t
\]

and by identification with \((11)\) the matrices

\[
A_t = 0_{3,3}, \quad Q(\hat{R}_t) = \hat{R}_t \text{Cov}(w_t) \hat{R}_t^T.
\]

In the same way, discarding terms of magnitude \( \|\xi_t\|\|\xi_t\| \), and \( \|\xi_t\|\|V_n\| \), and using the subsequent approximations \( \eta_t^+ = I_3 + (\xi_t)_x \), \( \exp(u) = I_3 + (u)_x \), \( \eta_t^{-1} = I_3 + (\xi_t)_x \) and \( \exp^{-1}(I_3 + (u)_x) = u, \) we get

\[
\xi_t^+ = \xi_t - L_n \begin{bmatrix} (g)_x, \xi_t^+ - \hat{R}_t V_n^g \hat{R}_n \end{bmatrix}
\]

yielding the matrices for the linearized update error equation \((12)\)

\[
H = \begin{bmatrix} (g)_x \end{bmatrix}^T, \quad N(\hat{R}_t) = \hat{R}_n \text{Cov}(V_n) \hat{R}_n^T
\]

The gains \( L_n \) are thus computed using the Riccati equation \((13)\), where the matrices \( A_t, H, Q(\hat{R}_t), N(\hat{R}_t) \) are defined by \((17)\) and \((18)\).

**Stability results** The IEKF, when used as an observer, enjoys the following desirable convergence property:

**Proposition 2.** Assume that the eigenvalues of \( \text{Cov}(V_n) \), \( \text{Cov}(w_t) \) are (uniformly) lower bounded by some \( \alpha > 0 \). Then the IEKF, used as an attitude observer for the continuous time system \( \frac{d}{dt} R_t = R_t (\omega_t)_x \) with an infinite number of discrete observations \( Y_n = (R_{t_n}^T g; R_{t_n}^T b) \) where \( g, b \) are known non-collinear vectors of \( \mathbb{R}^3 \), is such that \( \hat{R}_t R_t^T \to I_3 \) when \( t \to \infty \) for \( R_0, R_0 \) sufficiently close.

**Proof.** It suffices to verify the assumptions of the theorem, which is particularly simple as \( A_t \equiv 0 \). Using the latter equality and the assumptions on the noise covariance matrices readily implies assumptions \((1), (2), (3), (4)\) hold, and using moreover that \( g, b \) are not collinear, the matrix \( H \) is of rank \( 3 \) so that assumption \((5)\) holds. The infinite
number of observations is implicit in the theorem, it is explicitly mentioned here to underline to the practitioner the existence of this assumption.

4. SECOND EXAMPLE: GYROCOMPASS

This example is also well-known. Note that, in the recent paper Batista et al. (2014), the authors have obtained a globally converging observer for it in continuous time. The filter proposed in the present paper is only locally convergent, but it has the merit to strictly be an EKF variant, with all the associated advantages we have already largely discussed.

Model Consider the attitude of a gyrocompass, represented by the rotation matrix \( R_t \in SO(3) \) mapping the coordinates of a vector expressed in the gyrocompass frame to its coordinates in the static frame. The vehicle is endowed with gyrometers giving an angular velocity \( \omega_t \), and supposed precise enough to measure the earth rotation. The equation of the dynamics then reads:

\[
\frac{d}{dt} R_t = (\Omega) \times R_t + R_t (\omega_t + w_t) \times
\]

where \( \Omega \) is the earth rotation axis known in a inertial frame linked to remote stars, and \( w_t \) is the gyro noise. The gyrocompass being at rest, an accelerometer measures the acceleration of a vector expressed in the gyrocompass frame (quasi-static hypothesis):

\[
Y_n = R_t^T \omega \quad \text{and} \quad V_n^g
\]

The reader can verify the relation (2).

Error equation Due to similarities with the latter example, we skip quite a number of details in the sequel. The RIEKF (8)-(9) for the system (19)-(20) is defined by:

\[
\frac{d}{dt} \hat{R}_t = (\Omega) \times \hat{R}_t + \hat{R}_t (\omega_t + w_t) \times, \quad \hat{R}_{t_n} = \exp(L_n[\hat{R}_{t_n} g - g])\hat{R}_{t_n}
\]

The invariant error is \( \eta_t = \hat{R}_t R_t^T \) and its evolution reads:

\[
\frac{d}{dt} \eta_t = (\Omega) \times \eta_t - \eta_t (\Omega) \times - (\hat{R}_t w_t) \times \eta_t
\]

\[
\eta_{t_n}^+ = \exp(L_n[\eta_t g - g + \hat{R}_t V_{n}^g])\eta_{t_n}
\]

Linearized error equation and gain tuning To linearize this equation we introduce the linearized error \( \xi_t \) defined as \( \eta_t = I_3 + (\xi_t) \times \). Introducing \( \eta_t = I_3 + (\xi_t) \times, \eta_t^+ = I_3 + (\xi_t^+) \times, \exp(u) = I_3 + (u) \times \), and removing the second-order terms in \( \xi_t, V_{n}^g, w_t \) and products of those terms, we obtain the following linearized error equation:

\[
\frac{d}{dt} \xi_t = (\Omega) \times \xi_t - \hat{R}_t w_t, \quad \xi_{t_n}^+ = \xi_{t_n} - L_n[(\xi_t) \times + \hat{R}_t V_{n}^g]
\]

The gains \( L_n \) are thus computed using the Riccati equation (13) with \( A_t = (\Omega) \times, H = -(\xi_t) \times, Q(\hat{R}_t) = \hat{R}_t \text{Cov}(w_t)\hat{R}_t^T, N(\hat{R}_{t_n}) = \hat{R}_t \text{Cov}(V_n)\hat{R}_t^T \).

Stability results We start with a simple lemma.

Lemma 3. The system (19), (20) defining the gyrocompass is observable if \( g \) and \( \Omega \) are non-collinear.

Proof. The observability matrix associated to the deterministic part of (19), (20) for two updates reads:

\[
\left( \begin{array}{c} (g) \times \ \ (g) \times \\ \ (g) \times \ \ (g) \times \\ \end{array} \right), \quad \text{where} \quad R_O = \exp(\Omega(t_{n+1} - t_n)) \quad \text{denotes the rotation matrix corresponding to the axis-angle representation} \ \Omega.
\]

The rank is not increased if this matrix is multiplied on the left by \( \left( \begin{array}{c} I_3 \ 0_{3 \times 3} \end{array} \right) \). We obtain

\[
\left( \begin{array}{c} (g) \times \\ (g) \times \\ (g) \times \\ \end{array} \right) \quad \text{This is the same observability matrix as in the case where vectors} \ g \quad \text{and} \quad \Omega \quad \text{are non-collinear, i.e. if} \ g \quad \text{and} \quad \Omega \quad \text{are non-collinear.}
\]

Proposition 4. The IEKF for the gyrocompass problem (19)-(20) is an asymptotically stable observer if \( g \) and \( \Omega \) are non-collinear, and the eigenvalues of \( \text{Cov}(w_t), \text{Cov}(V_n) \) are greater than some \( \alpha > 0 \).

This is an immediate consequence of Lemma 3 and Theorem 1. Note that, if \( g \) and \( \Omega \) are collinear (in other word the gyrocompass is on the north or south pole), the heading is anyway not observable.

5. THIRD EXAMPLE: 3D CAMERA LOCALIZATION

This example has been partly considered in Hervier et al. (2012), except that here we do not assume to have the gyroscopes. No proof of convergence has been derived whatsoever in prior work.

Model We consider here the problem of a hand-held depth-camera (such as the Microsoft Kinect sensor) in a fixed environment of which we already possess a 3D model. The camera is supposed not to be equipped with motion sensors (such as gyrometers and accelerometers). A classical way to account for the latter, is to assume the camera to be fixed (low dynamics motion prior) and to let the filter know the motion model is uncertain. Mathematically this boils down to assume the camera follows a Brownian motion. The attitude is denoted by the rotation matrix \( R_t \) and the position by the 3-dimensional vector \( x_t \). The equations then read:

\[
\frac{d}{dt} R_t = (w^R_t) \times R_t, \quad \frac{d}{dt} x_t = (w^R_t) \times x + w^x_t
\]

where \( w^R_t \) and \( w^x_t \) are noises. The process noise is assumed known in the static frame, as no matter what the camera
The right-invariant error is \( \eta \). The system must be embedded into a matrix Lie group in order to apply the theory above. The chosen matrix Lie group is \( SE(3) \) (see Appendix A.2), and the embedding is done through the following alternative state, output, and noise (matrix) variables:

\[
\begin{align*}
\chi_t &= \begin{pmatrix} R_t & x_t \\ 0_{1 \times 3} & 1 \end{pmatrix}, & w_t &= \begin{pmatrix} (w^R_t) \times w^x_t \\ 0_{1 \times 3} \end{pmatrix} \\
Y_n &= \begin{pmatrix} Y_n^R - Y_n^L \\ 0_{1 \times 3} \end{pmatrix}, & V_n &= \begin{pmatrix} V_n^R - V_n^L \\ 0_{1 \times 3} \end{pmatrix}
\end{align*}
\]

Equations (21), (22) then become:

\[
\begin{align*}
\frac{d}{dt} \chi_t &= w_t \chi_t & (24) \\
Y_n &= \chi_n^{-1} V_n & (25)
\end{align*}
\]

Error equation Note that, the output here is slightly different from (7) as \( V_n \) (and thus \( Y_n \)) is here a matrix. But the IEKF theory can be easily generalized to this case. Indeed, the IEKF equations are unchanged, apart from the linearization process that is slightly modified as explained in the sequel. It is easily seen the theorem still holds (as proved in the first version of the preprint of Barrau and Bonnabel (2017) available on Arxiv). The RIEKF (8)-(9) for the system (24)-(25) is:

\[
\frac{d}{dt} \chi_t = 0_{4,4}, \quad \chi_n^+ = \exp(L_n[\exp^{-1}(\chi_n Y_n)]) \chi_n
\]

The right-invariant error is \( \eta = \chi_n^{-1}. \) Using the general equality \( \frac{d}{dt} \chi_n^{-1} = -\chi_n^{-1} \frac{d}{dt} \chi_n \chi_n^{-1} \) its evolution reads:

\[
\frac{d}{dt} \eta = -\eta w_t, \quad \eta_n^+ = \exp(L_n[\exp^{-1}(\eta_n V_n)]) \eta_n
\]

Linearized error equation and gain tuning To linearize this equation we introduce the linearized error \( \xi_t \) through the following approximation \( \eta = I_4 + L_{sec(3)}(\xi). \) The observation noise now lies on the group and not in a vector space. The IEKF gain acts on the Lie algebra of the group, so that the noise must be defined in the Lie algebra as well. To do so, we assume it be not too large and use the following approximation \( V_n = I_4 + L_{sec(3)}(\eta_n). \)

Introducing \( \eta_{\xi}^+ = I_4 + L_{sec(3)}(\xi^+), \) \( \exp(u) = I_4 + L_{sec(3)}(u), \) \( \eta_{\xi}^{-1} = I_4 - L_{sec(3)}(\xi), \) \( \exp^{-1}(I_4 - L_{sec(3)}(u)) = u \) and the approximation \( \exp^{-1}(ab) = \exp^{-1}(a) + \exp^{-1}(b) \) for small \( a, b \) in (26), and removing the second-order terms in \( \xi_t, \) \( \exp^{-1}(V_n) \) and \( w_t \) we obtain:

\[
\frac{d}{dt} \xi_t = -w_t, \quad \xi_t^+ = \xi_t - L_n(\xi_t - \eta_n)
\]

The gains \( L_n \) are computed using the Riccati equation (13) where the matrices \( A_t, H, Q(\chi_t) \) and \( N(\chi_t) \) are defined according to the linearized equation above by \( A_t = 0_{3,3}, \) \( H = -I_6, \) and \( Q(\chi_t) = \text{Cov}(w_t), \) \( N(\chi_t) = \text{Cov}(v_n) = \text{Cov}(\exp^{-1}(V_n)) \)

The stability properties of the IEKF imply (see the remark above about the validity of the theorem in that case):

**Proposition 5.** The IEKF for the movement estimation problem (24), (25) is an asymptotically stable observer as long as all the eigenvalues of \( \text{Cov}(\exp^{-1}(V_n)) \), and at least one eigenvalue of \( \text{Cov}(w_t) \) are lower bounded by some fixed \( \alpha > 0 \).
Appendix A. LIE GROUPS BACKGROUND

A matrix Lie group \( G \) is a subset of square invertible \( N \times N \) matrices \( \mathcal{M}_N(\mathbb{R}) \) verifying the following properties:

\[
\begin{align*}
Id & \in G, & \forall g \in G, g^{-1} \in G, & \forall a, b \in G, ab \in G \\
\end{align*}
\]

If \( \gamma(t) \) is a curve over \( G \) with \( \gamma(0) = Id \), then its derivative at \( t = 0 \) necessarily lies in a subset \( \mathfrak{g} \) of \( \mathcal{M}_N(\mathbb{R}) \). \( \mathfrak{g} \) is a vector space and it is called the Lie algebra of \( G \) and has same dimension as \( G \). Thanks to a linear map denoted by \( \mathcal{L}_g : \mathbb{R}^{\dim \mathfrak{g}} \rightarrow \mathfrak{g} \), one can advantageously identify \( \mathfrak{g} \) to \( \mathbb{R}^q \) where \( q = \dim \mathfrak{g} \). Besides, the vector space \( \mathfrak{g} \) can be mapped to the matrix Lie group \( G \) through the classical matrix exponential \( \exp_m \). Thus, \( \mathbb{R}^q \) can be mapped to \( G \) through the Lie exponential map defined by \( \exp(\xi) := \exp_m(\mathcal{L}_0(\xi)) \) for \( \xi \in \mathbb{R}^q \). We have thus \( \exp(\xi) = Id + \mathcal{L}_g(\xi) + O(\xi^2) \in \mathcal{M}_N(\mathbb{R}) \). For all \( \xi \in \mathbb{R}^q \) the adjoint matrix \( ad_G \in \mathbb{R}^{q \times q} \) is very useful to linearize equations over Lie groups. It satisfies \( ad_G \xi = -ad_G \xi \) and

\[
\mathcal{L}_g(\xi) \mathcal{L}_g(\zeta) - \mathcal{L}_g(\xi) \mathcal{L}_g(\zeta) = \mathcal{L}_g(-ad_G \xi) \tag{A.1}
\]

A.1 Group of rotation matrices \( SO(3) \)

We have here \( G = \mathcal{SO}(3) = \{ R \in \mathcal{M}_3(\mathbb{R}), R R^T = Id \} \). The tangent space around identity is \( \mathfrak{so}(3) = \{ A \in \mathcal{M}_3(\mathbb{R}), A = -A^T \} \), the space of skew-symmetric matrices. An isomorphism between \( \mathbb{R}^3 \) and \( \mathfrak{g} \) is given by \( L_{\mathfrak{so}(3)}(\xi_1) = (\xi_1) = \left( \begin{array}{ccc} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{array} \right) \). The exponential mapping is given by the formula:

\[
\exp(\xi) = I_3 + \frac{\sin(||\xi||)}{||\xi||} \left( \begin{array}{ccc} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{array} \right) + \frac{2\sin(||\xi||/2)^2}{(||\xi||)^2} \left( \begin{array}{ccc} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{array} \right).
\]

A.2 Group of direct spatial isometries \( SE(3) \)

We have here \( G = SE(3) = \{ \left( \begin{array}{cc} R & x \\ 0_{1 \times 3} & 1 \end{array} \right), R \in SO(3), x \in \mathbb{R}^3 \} \). The tangent space around identity is \( \mathfrak{se}(3) = \{ (\xi \times x, \xi), x \in \mathbb{R}^3 \} \). An isomorphism between \( \mathbb{R}^6 \) and \( \mathfrak{se}(3) \) is given by \( L_{\mathfrak{se}(3)}(\xi x) = (\xi \times x, \xi) \). The exponential mapping is given by the formula:

\[
\exp(\xi x) = I_4 + S + \frac{1 - \cos(||\xi||)}{||\xi||^2} S^2 + \frac{||\xi|| \sin(||\xi||)}{||\xi||^3} S^3, \text{ where } S = L_{\mathfrak{se}(3)}(\xi x).
\]