A Generalized Popov–Belevitch-Hautus Test of Observability

Bijoy K. Ghosh and Joachim Rosenthal

***Abstract*—In this paper, an earlier result on the problem of observability of a linear dynamical system due to Popov–Belevitch–Hautus has been generalized and applied to the problem of observing the initial condition of a linear dynamical system described on the space of d dimensional affine planes in $\mathbb{R}^n$.**

I. Introduction and Motivation

In this paper we generalize the well known Popov-Belevitch-Hautus test (see [3]) on the observability of a linear dynamical system. Let $\mathbb{K}$ denote either the field of real ($\mathbb{K} = \mathbb{R}$) or the field of complex ($\mathbb{K} = \mathbb{C}$) numbers. Let A be an $n \times n$ matrix and let C be a $p \times n$ matrix defined over $\mathbb{K}$. Consider the linear time invariant system

$$\dot{x} = Ax, \qquad y = Cx, \qquad x \in \mathbb{K}^n, y \in \mathbb{K}^p. \tag{1.1}$$

The well-known Hautus test [3] gives a necessary and sufficient condition, when the state vector $x(t)$ can be observed from the output measurement $y(t)$. To be precise one has the following.

Theorem 1 (Hautus [3]): System (1.1) is observable over either $\mathbb{R}$ or $\mathbb{C}$ if, and only if

$$\text{rank}\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n, \quad \text{for all } \lambda \in \mathbb{C}. \tag{1.2}$$

It may be remarked that the rank can only be less than n if λ is an eigenvalue of the matrix A.

In this paper we consider dynamical systems for which the output vector is not observed exactly but can be ascertained with an ambiguity restricted to a $d-$dimensional affine subsapce. The problem that we propose to consider is to compute if possible the initial condition and hence the states of the dynamical system up to a $d-$dimensional affine subspace. Thus for the dynamical system (1.1), if we assume that the output vector $y(t)$ is observed up to a $d-$dimensional plane given by an equation of the form

$$\theta(t)y(t) = \eta(t) \tag{1.3}$$

where $\theta(t)$ is a $(p-d) \times n$ matrix function of time having full rank for all but countably many instances of time and $\eta(t)$ is a vector function of time. The problem is to derive conditions on A and C under which $x(0)$ can be observed up to a $d-$dimensional plane.

The above class of problem occurs in machine vision as has already been introduced in [6], [1]. Specifically if we consider a plane in $\mathbb{R}^3$ with coordinates (X, Y, Z) given by

$$sZ = pX + qY + r. \tag{1.4}$$

Manuscript received October 25, 1993; revised March 8, 1994 and May 14, 1994. B. K. Ghosh was supported in part by the Department of Energy under Grant DE-FG-02-90ER14140. J. Rosenthal was supported in part by the National Science Foundation under Grant DMS-9201263.

B. K. Ghosh is with the Department of Systems Science and Mathematics, Washington University, St. Louis, MO 63130 USA.

J. Rosenthal is with the Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556 USA.

IEEE Log Number 9406983.

Let us assume that the points on the plane (1.4) satisfy a dynamical system

$$\dot{\chi} = A\chi + b \tag{1.5}$$

where A is an arbitrary 3×3 matrix and b is a 3×1 vector given by

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{1.6}$$

$$b = (b_1 \quad b_2 \quad b_3)^T, \tag{1.7}$$

respectively, and where χ is given by

$$\chi = (X \quad Y \quad Z)^T.$$

One can compute a dynamical system for the shape parameters p, q, r, and s described as follows:

$$\frac{d}{dt}\begin{pmatrix} p \\ q \\ -s \\ r \end{pmatrix} = \begin{pmatrix} -A^T & 0 \\ -b^T & 0 \end{pmatrix}\begin{pmatrix} p \\ q \\ -s \\ r \end{pmatrix}. \tag{1.8}$$

Typically a point on the plane (1.4) is observed with the aid of a CCD camera that projects (X, Y, Z) perspectively onto an image plane. Let (η_1, η_2) be the coordinates of the image plane and assume that the perspective projection is defined as

$$\eta_1 = \frac{fX}{Z+f}, \qquad \eta_2 = \frac{fY}{Z+f} \tag{1.9}$$

where f is the focal length of the camera. One can easily compute a differential equation that is satisfied by the coordinates (η_1, η_2) and is given by

$$\begin{aligned} y_9\dot{\eta}_1 &= y_1 + y_3\eta_1 + y_4\eta_2 + \frac{1}{f}(y_7\eta_1^2 + y_8\eta_1\eta_2) \\ y_9\dot{\eta}_2 &= y_2 + y_6\eta_2 + y_5\eta_1 + \frac{1}{f}(y_8\eta_2^2 + y_7\eta_1\eta_2) \end{aligned} \tag{1.10}$$

where

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -fb_1 & fa_{13} \\ 0 & 0 & -fb_2 & fa_{23} \\ -b_1' & 0 & b_3 - fa_{11} & a_{11} - a_{33} \\ 0 & -b_1' & -fa_{12} & a_{12} \\ -b_2' & 0 & -fa_{21} & a_{21} \\ 0 & -b_2' & b_3 - fa_{22} & a_{22} - a_{33} \\ -b_3' & 0 & -fa_{31} & a_{31} \\ 0 & -b_3' & -fa_{32} & a_{32} \\ 0 & 0 & -f & 1 \end{pmatrix}\begin{pmatrix} p \\ q \\ -s \\ r \end{pmatrix} \tag{1.11}$$

and where

$$b' = (b_1 - a_{13}f,\ b_2 - a_{23}f,\ b_3 - a_{33}f) \triangleq (b_1', b_2', b_3'). \tag{1.12}$$

The equation (1.10) is known as the "optical flow" and in the literature various algorithms exist as to how one can estimate $(\dot{\eta}_1, \dot{\eta}_2)$ for a given pair (η_1, η_2) (see Horn [4]).

For our purposes we would like to view (1.8), (1.11) as a linear system for which the output vector y_j is not observed but instead one observes the vector $(\dot{\eta}_1, \dot{\eta}_2, \eta_1, \eta_2)$ at various points on the image plane. Note that for almost every point on the image plane, (1.10) describes a homogeneous seven-dimensional plane in $\mathbb{R}^9$. Thus if one observes $(\dot{\eta}_1, \dot{\eta}_2, \eta_1, \eta_2)$ for 3 points on the image plane, the output vector in (1.11) is observed up to a homogeneous 3-plane.

On the other hand, if 4 points are observed the output vector in (1.11) is observed up to a homogeneous line. Various other cases can be demonstrated likewise. Note in particular that by observing the vector $(\dot{\eta}_1, \dot{\eta}_2, \eta_1, \eta_2)$ for 1 or 2 points on the image plane, the output vector is observed up to respectively a seven- or five-dimensional plane in $\mathbb{R}^9$. Such an observation does not shed any new information on the vector (p, q, s, r). In practice if the vector (p, q, s, r) is recovered only up to a d-dimensional plane where $d > 1$, one would typically use multiple cameras to determine the exact value of (p, q, s, r).

II. Problem Formulation and Main Result

In order to introduce the main result considered in this paper, let $P_0 \subset \mathbb{K}^n$ be a d-dimensional affine subspace not necessarily passing through the origin. In this paper we shall use the expression "d-dimensional affine subspace" to mean a "d-dimensional plane." We say that the dynamical system (1.1) observes P_0 if for any $0 \le t_1 < t_2$ it is possible to calculate P_0 from the observation of the "moving plane" $CP(t) \triangleq Ce^{At}P_0$ in $\mathbb{K}^p, t_1 \le t \le t_2$. Our main theorem is described as follows.

Theorem 2 (Main Theorem): System (1.1) observes any d-dimensional affine subspace P_0 in $\mathbb{K}^n$ if for any set of eigenvalues $\lambda_0, \cdots, \lambda_d$ of A one has

$$\operatorname{rank}\begin{bmatrix}(A-\lambda_0 I)\cdots(A-\lambda_d I)\\ C\end{bmatrix} = n. \tag{2.1}$$

Moreover this condition is also necessary if $d = 0$ or if all eigenvalues of the matrix A are in $\mathbb{K}$.

Remark 3: Note that over the complex numbers $\mathbb{C}$, condition (2.1) is necessary and sufficient. Moreover if $d = 0$, Theorem 2 reduces to Theorem 1. Finally if $d = 1$ this result implies the one given in [6] due to Wang, Martin, Dayawansa, and Ghosh.

The following two examples explain the ingredients of our result.

Example 4: Consider the real system

$$\dot{x} = Ax = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}x, \qquad y = Cx = (1 \quad 0)x, \qquad x \in \mathbb{R}^2. \tag{2.2}$$

Because the eigenvalues of A are real, condition (2.1) in Theorem 2 is necessary and sufficient. In particular if $p_0 \in \mathbb{R}^2$ is a point it can be observed from $y(t) = Ce^{At}p_0$ because

$$\operatorname{rank}\begin{bmatrix}-\lambda & 1\\ 0 & -\lambda\\ 1 & 0\end{bmatrix} = 2 \tag{2.3}$$

for all $\lambda \in \mathbb{R}$ including the case when λ is an eigenvalue of A. However if $l_0 \subset \mathbb{R}^2$ is a line, it cannot be observed from $l(t) = Ce^{At}l_0$ because for every pair of lines l_0 and l_1 in $\mathbb{R}^2$ and for all but a finite number of time instants t, we have

$$\{\xi : \xi = Ce^{At}\delta,\ \delta \in l_0\} = \{\xi : \xi = Ce^{At}\delta,\ \delta \in l_1\}. \tag{2.4}$$

Thus the lines l_0 and l_1 are both mapped to the entire real axis and therefore they cannot be observed. We also note that

$$\operatorname{rank}\begin{bmatrix}\lambda_0\lambda_1 & -\lambda_0-\lambda_1\\ 0 & \lambda_0\lambda_1\\ 1 & 0\end{bmatrix} \neq 2 \quad \text{for every pair of eigenvalues } \lambda_0, \lambda_1 \text{ of } A. \tag{2.5}$$

In fact for $\lambda_0 = \lambda_1 = 0$, rank drops to 1.

Remark 5: Note that the equality of the two sets in (2.4) is valid for all but possibly a finite number of time instants. One might correctly conclude from this, that in principle observability can be ascertained on the basis of these finitely many exception points. However we would still like to say that the lines l_0 and l_1 are unobservable on the basis of any arbitrary time interval (t_1, t_2).

Example 6: Consider the fourth order system

$$\dot{x} = Ax = \begin{pmatrix}-81 & -56 & 57 & -11\\ 146 & 102 & -106 & 20\\ 62 & 43 & -46 & 9\\ 203 & 138 & -149 & 31\end{pmatrix}x,$$
$$y = Cx = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}x, \qquad x \in \mathbb{R}^4. \tag{2.6}$$

A direct computation shows that the pair (A, C) is observable and the matrix A has eigenvalues $0, 1, 2, 3$. Since for any 2 eigenvalues λ_0, λ_1 the nullspace of $(A-\lambda_0 I)(A-\lambda_1 I)$ is equal to the sum of the eigenspaces $\operatorname{Ker}(A-\lambda_0 I)$ and $\operatorname{Ker}(A-\lambda_1 I)$ and none of those sums contains the vector $(0, 0, 0, 1)^T \in \operatorname{Ker}(C)$ it follows from Theorem 2 that if $l_0 \subset \mathbb{R}^4$ is a line in $\mathbb{R}^4$, it can be observed from $Cl(t) = Ce^{At}l_0$, which is a motion of lines in $\mathbb{R}^3$. On the other hand one immediately verifies that

$$A(A-I)(A-2I) = \begin{pmatrix}-3 & -3 & 3 & 0\\ 8 & 8 & -8 & 0\\ -1 & -1 & 1 & 0\\ -23 & -23 & 23 & 0\end{pmatrix}.$$

It therefore follows that certain two-planes $P_0 \subset \mathbb{R}^4$ cannot be observed from $Ce^{At}P_0 \subset \mathbb{R}^3$. Specifically consider the vectors $v_0 := (-12, 20, 8, 28)^T, v_1 := (35, -60, -25, -85)^T$, and $v_2 := (-23, 40, 17, 58)^T$. One immediately verifies that v_0, v_1, v_2 are eigenvectors corresponding to the eigenvalues 0, 1, and 2. Also note that $v_0 + v_1 + v_2 = (0, 0, 0, 1)^T$. Let P be the three-dimensional subspace in $\mathbb{R}^4$ spanned by the vectors v_0, v_1, and v_2. It can be verified that for all but a finite set of values of $t, Ce^{At}P$ is a two-dimensional plane in $\mathbb{R}^3$. To see this, note that $Ce^{At}v_j = Ce^{\lambda_j t}v_j$ for $j = 0, 1, 2$ where λ_j is the eigenvalue corresponding to the eigenvector v_j. Since Cv_0, Cv_1, Cv_2 are linearly dependent, it follows that $Ce^{At}v_0, Ce^{At}v_1, Ce^{At}v_2$ are linearly dependent as well. Thus for any $0 \le t_1 < t_2$ and for almost every pair of two-dimensional planes Q_1 and Q_2 such that $Q_1 \neq Q_2$ and $Q_1 \subset P, Q_2 \subset P$, we have

$$Ce^{At}Q_1 = Ce^{At}Q_2$$

for $t_1 \le t \le t_2$. Hence the planes Q_1 and Q_2 cannot be observed.

III. An Associated Dynamical System

The proof of Theorem 2 will be broken down in a sequence of lemmas. The proof is mainly based on a careful study of a dynamical system defined on the $\mathbb{K}$ vector space $\bigwedge^k \mathbb{K}^n$, the k-fold wedge product of $\mathbb{K}^n$ (see [2] for a reference). This system has also been used in [6] to derive the results there.

First recall that $\bigwedge^k \mathbb{K}^n$ is linearly generated by the vectors

$$\{x_1 \wedge \cdots \wedge x_k | x_i \in \mathbb{K}^n,\ i = 1, \cdots, k\}.$$

Addition in $\bigwedge^k \mathbb{K}^n$ is multilinear and alternating in the components. If $\{e_1, \cdots, e_n\}$ is a basis of $\mathbb{K}^n$ then it follows from the multilinearity and the alternating property of the wedge product that

$$\mathcal{B} := \{e_{i_1} \wedge \cdots \wedge e_{i_k} | 1 \le i_1 < \cdots < i_k \le n\}$$

is a basis of $\bigwedge^k \mathbb{K}^n$. In particular one has $\dim \bigwedge^k \mathbb{K}^n = \binom{n}{k}$. If a vector $v \in \bigwedge^k \mathbb{K}^n$ has a representation $v = x_1 \wedge \cdots \wedge x_k$ for some particular vectors $x_i \in \mathbb{K}^n$, $i = 1, \cdots, k$, one says that v is a *decomposable* vector. The coordinates of a decomposable vector with respect to the canonical basis $\mathcal{B}$ are sometimes called the Plücker coordinates of v.

Next define linear maps

$$\hat{A}: \bigwedge^k \mathbb{K}^n \to \bigwedge^k \mathbb{K}^n$$
$$x_1 \wedge \cdots \wedge x_k \mapsto \sum_{i=1}^{k} x_1 \wedge \cdots \wedge x_{i-1} \wedge Ax_i \wedge x_{i+1} \cdots \wedge x_k \tag{3.1}$$

and

$$\hat{C}: \bigwedge^k \mathbb{K}^n \to \bigwedge^k \mathbb{K}^p$$
$$x_1 \wedge \cdots \wedge x_k \mapsto Cx_1 \wedge \cdots \wedge Cx_k. \tag{3.2}$$

$\hat{A}$ and $\hat{C}$ induce the dynamical system

$$\dot{\hat{X}} = \hat{A}\hat{X}, \qquad \hat{Y} = \hat{C}\hat{X}. \tag{3.3}$$

The state space of (3.3) is the vector space $\bigwedge^k \mathbb{K}^n$ and the output space is the vector space $\bigwedge^k \mathbb{K}^p$. We would like to remark that if the trajectories $Cx_1(t), \cdots, Cx_k(t)$ are solutions of the system (1.1) then $\hat{C}(x_1(t) \wedge \cdots \wedge x_k(t))$ is a solution of system (3.3). It is our goal to show that, provided the eigenvalues of A are in $\mathbb{K}$, (2.1) is equivalent to a particular notion of observability of the system (3.3) and that this condition is also necessary and sufficient for the observability of P_0 under the output function $Ce^{At}P_0$. The following lemmas prepare for this result.

Lemma 7: The (unique) solution of the initial value problem

$$\frac{d}{dt}(x_1(t) \wedge \cdots \wedge x_k(t)) = \hat{A}(x_1(t) \wedge \cdots \wedge x_k(t))$$
$$x_1(0) \wedge \cdots \wedge x_k(0) = v_1 \wedge \cdots \wedge v_k$$

is given through

$$x_1(t) \wedge \cdots \wedge x_k(t) = e^{At}v_1 \wedge \cdots \wedge e^{At}v_k. \tag{3.4}$$

Proof: Differentiate (3.4) and recall the definition of $\hat{A}$. Q.E.D.

Lemma 8: Let $x_1, x_2, \cdots, x_k \in \mathbb{K}^n$ be vectors and $c_1, c_2, \cdots, c_k \in \mathbb{K}$ be scalars. Let $c \triangleq c_1 + c_2 + \cdots + c_k$. Then it follows that

$$(\hat{A} - cI)(x_1 \wedge x_2 \wedge \cdots \wedge x_k) = \sum_{i=1}^{k} x_1 \wedge \cdots \wedge x_{i-1} \wedge (A - c_iI)x_i \wedge x_{i+1} \cdots \wedge x_k. \tag{3.5}$$

Proof: Note that

$$\begin{aligned}(\hat{A} - cI)(x_1 \wedge x_2 \wedge \cdots \wedge x_k) &= (Ax_1 \wedge x_2 \wedge \cdots \wedge x_k) \\ &\quad - c_1(x_1 \wedge x_2 \cdots \wedge x_k) \\ &\quad + (x_1 \wedge Ax_2 \wedge \cdots \wedge x_k) \\ &\quad - c_2(x_1 \wedge x_2 \cdots \wedge x_k) \\ &\quad + \cdots \\ &\quad \cdot \;\cdots \\ (x_1 \wedge x_2 \wedge \cdots \wedge Ax_k) - c_k(x_1 \wedge x_2 \cdots \wedge x_k) &= \sum_{i=1}^{k} x_1 \wedge \cdots \wedge x_{i-1} \wedge (A - c_iI) \\ &\quad \cdot x_i \wedge x_{i+1} \cdots \wedge x_k.\end{aligned}$$

Q.E.D.

Lemma 9: Let $\{x_1, \cdots, x_n\} \subset \mathbb{K}^n$ be a $\mathbb{K}$-basis of generalized eigenvectors of the matrix A having corresponding eigenvalues $\{\lambda_1, \cdots, \lambda_n\}$ (possibly repeated) then

$$\{x_{i_1} \wedge \cdots \wedge x_{i_k} | 1 \le i_1 < \cdots < i_k \le n\} \tag{3.6}$$

is a $\mathbb{K}$-basis of generalized eigenvectors of the matrix $\hat{A}$ with corresponding eigenvalues $\lambda_{i_1} + \cdots + \lambda_{i_k}$.

Proof: Clearly the set of vectors (3.6) are linearly independent and therefore form a basis. Assume that the vectors $x_{i_1} \cdots x_{i_k}$ have a nilpotency index $m_{i_1} \cdots m_{i_k}$, i.e.,

$$(A - \lambda_{i_k}I)^{m_{i_k}-1}x_{i_k} \neq 0, \qquad (A - \lambda_{i_k}I)^{m_{i_k}}x_{i_k} = 0. \tag{3.7}$$

In particular, if $m_{i_k} = 1$ it follows that x_{i_k} is an eigenvector with λ_{i_k} being the corresponding eigenvalue. Let us define

$$q = m_{i_1} + \cdots + m_{i_k} + 1 - k \tag{3.8}$$

it is trivial to verify using Lemma 8 that

$$(\hat{A} - (\lambda_{i_1} + \cdots + \lambda_{i_k})I)^q x_{i_1} \wedge \cdots \wedge x_{i_k} = 0. \tag{3.9}$$

Q.E.D.

Lemma 10: Let N be a nilpotent operator acting on $\mathbb{K}^q$. For every vector $v \in \mathbb{K}^q$ there is a unique $u \in \mathbb{K}^q$ such that

$$v = u + Nu + \cdots + N^q u.$$

Moreover if m is the nilpotency index of v then $\{u, \cdots, N^{m-1}u\}$ are linearly independent.

Proof: The unique vector u is given through $u := (I - N)v$. The linear independence is clear. Q.E.D.

Before we state the next result we note the following.

Remark 11: Note that not every vector in the vector space $\bigwedge^k \mathbb{K}^n$ is of the form $x_1 \wedge \cdots \wedge x_k$ and those that are, would be known as decomposable vectors.

The next result establishes the crucial relation between the observability of the pair $(\hat{A}, \hat{C})$ and the condition (2.1).

Proposition 12: Assume that the eigenvalues of the matrix A are in $\mathbb{K}$. Then the following conditions are equivalent:

1)There are eigenvalues $\lambda_{i_1}, \cdots, \lambda_{i_k}$ of A and a nonzero vector $v \in \mathbb{K}^n$ such that

$$\begin{pmatrix}(A - \lambda_{i_1}I) \cdots (A - \lambda_{i_k}I) \\ C\end{pmatrix} v = 0. \tag{3.10}$$

2)There is a $\lambda \in \mathbb{K}$ and a decomposable vector $\beta_1 \wedge \cdots \wedge \beta_k \in \bigwedge^k \mathbb{K}^n$ such that

$$\begin{pmatrix}\hat{A} - \lambda I \\ \hat{C}\end{pmatrix} \beta_1 \wedge \cdots \wedge \beta_k = 0. \tag{3.11}$$

3)The dynamical system (3.3) has a decomposable vector $\alpha_1 \wedge \cdots \wedge \alpha_k \in \bigwedge^k \mathbb{K}^n$ in its unobservable subspace.

Proof: 1)→2): Let $\lambda_1, \cdots, \lambda_s$ be the eigenvalues of A and let

$$\mathbb{K}^n = \oplus_{i=1}^{s} V_{\lambda_i} \tag{3.12}$$

be the decomposition into generalized eigenspaces. This decomposition induces a decomposition

$$v = v_1 + \cdots + v_s.$$

Let $\Lambda = \{l_1, \cdots, l_p\}$ be the eigenvalues appearing in the product

$$P := (A - \lambda_{i_1}I) \cdots (A - \lambda_{i_k}I)$$

and denote by $m(l_1), \cdots, m(l_p)$ their multiplicity, i.e., we have

$$P = (A - l_1I)^{m(l_1)} \cdots (A - l_pI)^{m(l_p)}.$$

From the A invariance of the generalized eigenspaces V_{λ_i} it follows that $v_h = 0$ if $h \notin \Lambda$. Moreover if $h \in \Lambda$ then v_h has nilpotency index at most $m(h)$. In the following we restrict ourselves to the case when v_h has nilpotency index $m(h)$. The (easier) other cases are similar. By Lemma 10 we have an expansion

$$v = \sum_{j_1=0}^{m(l_1)-1} (A - l_1 I)^{j_1} u_{l_1} + \cdots + \sum_{j_p=0}^{m(l_p)-1} (A - l_p I)^{j_p} u_{l_p}. \quad (3.13)$$

In this summation there are $m(l_1) + \cdots + m(l_p) = k$ summands which we like to denote by $\beta_1, \cdots, \beta_k$. By Lemma 10 those vectors are linearly independent and from Lemma 9 it follows that $\beta := \beta_1 \wedge \cdots \wedge \beta_k$ is an eigenvector of $\hat{A}$ with corresponding eigenvalue $\lambda := \lambda_{i_1} + \cdots + \lambda_{i_k}$. Finally, from (3.10) it follows that $\{C\beta_1, \cdots, C\beta_k\}$ is a linearly dependent set. It follows

$$\begin{pmatrix} \hat{A} - (\lambda_{i_1} + \cdots + \lambda_{i_k})I \\ \hat{C} \end{pmatrix} \beta_1 \wedge \cdots \wedge \beta_k = 0. \quad (3.14)$$

2)→3): The vector $\beta_1 \wedge \cdots \wedge \beta_k$ is necessarily an eigenvector of $\hat{A}$ and therefore in the unobservable subspace U of the system (3.3).

3)→1): The fact that condition 3) implies condition 1) is nontrivial. Our proof follows mainly ideas already developed in [6] and in principle it should be possible to generalize the proof given in [6]. This however amounts to a large case by case search. In order to avoid those tedious arguments we will deviate at a crucial point from this program.

The proof is structured as follows. Consider the decomposable vector $\alpha := \alpha_1 \wedge \cdots \wedge \alpha_k$ in the unobservable subspace U of the system (3.3) whose existence we assume. Using the fact that U is $\hat{A}$ invariant we will construct a polynomial $f(x) \in \mathbb{K}[x]$ which has the property that $f(\hat{A})\alpha$ is a decomposable eigenvector of $\hat{A}.f(\hat{A})\alpha$ is then necessarily in the unobservable subspace U and this implies 2) and from there we will imply 1). The details are now described as follows.

Consider the set of eigenvalues $\{\lambda_1, \cdots, \lambda_s\} \subset \mathbb{K}$ of A and arrange the order such that

$$\operatorname{Re}\lambda_i < \operatorname{Re}\lambda_{i+1}$$

or

$$\operatorname{Re}\lambda_i = \operatorname{Re}\lambda_{i+1}$$

and

$$\operatorname{Im}\lambda_i < \operatorname{Im}\lambda_{i+1}.$$

Let us choose a set of generalized eigenvectors $\{x_1, \cdots, x_n\}$ of A and consider the decomposition of $\mathbb{K}^n$ into generalized eigenspaces given through $\mathbb{K}^n = \oplus_{i=1}^s V_{\lambda_i}$. Arrange the order of $\{x_1, \cdots, x_n\}$ in such a way that $\{x_1, \cdots, x_{i_1}\}$ forms a basis of V_{λ_1}, $\{x_{i_1+1}, \cdots, x_{i_2}\}$ forms a basis of V_{λ_2} and so on.

Let $\alpha := \alpha_1 \wedge \cdots \wedge \alpha_k$ be a decomposable vector in the unobservable subspace U. Expand $\alpha_j = \sum_{i=1}^n b_{ji} x_i$, $j = 1, \cdots, k$, in terms of this basis. In this way we associate to α a coefficient matrix $B = b_{ij}$ whose entries are unique up to premultiplication by an element of the special linear group $Sl_k := \{T \in Mat_{n \times n} | \det(T) = 1\}$. Without loss of generality we can therefore assume that the matrix B is in echelon form.

Consider now the decomposition of $\bigwedge^k \mathbb{K}^n$ into generalized eigenspaces.

$$\bigwedge^k \mathbb{K}^n = \oplus_{\lambda = \lambda_{i_1} + \cdots + \lambda_{i_k}} W_\lambda. \quad (3.15)$$

If w_λ denotes the component of $\alpha_1 \wedge \cdots \wedge \alpha_k$ in W_λ then

$$w_\lambda = \sum v_{i_1} \wedge \cdots \wedge v_{i_k}, \quad (3.16)$$

where v_{i_j} is the component of α_j in $V_{\lambda_{i_j}}$ and where the summation is taken over all indexes $(i_1, \cdots, i_k)$ having the property that $\lambda_{i_1} + \cdots + \lambda_{i_k} = \lambda$.

From the fact that the matrix B is assumed to be in echelon form and from the assumption that $\lambda_1, \cdots, \lambda_s$ are ordered, it follows that there is one eigenvalue μ such that the component w_μ of $\alpha_1 \wedge \cdots \wedge \alpha_k$ in W_μ is nonzero and decomposable, i.e.,

$$w_\mu = v_{r_1} \wedge \cdots \wedge v_{r_k}. \quad (3.17)$$

Indeed, v_{r_j} can be chosen in the following way. For $j = 1, \cdots, k$ consider the decomposition

$$\alpha_j = \sum_{i_j=1}^s v_{i_j} \quad (3.18)$$

induced by the eigenspace decomposition (3.12). Then choose r_j as the first index with the property, that $v_{r_j} \neq 0$. By definition $v_{r_1} \wedge \cdots \wedge v_{r_k}$ is nonzero, decomposable, and it represents the component of $\alpha_1 \wedge \cdots \wedge \alpha_k$ in W_μ.

Let m be the order of w_μ. It is our goal to calculate the eigenvector $(\hat{A} - \mu I)^{m-1} w_\mu$ and to show that this vector is decomposable as well. For this consider the initial value problem

$$\frac{d}{dt}(x_1(t) \wedge \cdots \wedge x_k(t)) = (\hat{A} - \mu I)(x_1(t) \wedge \cdots \wedge x_k(t))$$
$$x_1(0) \wedge \cdots \wedge x_k(0) = v_{r_1} \wedge \cdots \wedge v_{r_k}.$$

Using Lemma 7 and Lemma 8 one verifies that the solution is given through

$$e^{(\hat{A} - \mu I)t} w_\mu = e^{(A - \lambda_{r_1} I)t} v_{r_1} \wedge \cdots \wedge e^{(A - \lambda_{r_k} I)t} v_{r_k}. \quad (3.19)$$

Because $v_{r_j} \in V_{\lambda_{r_j}}$, $j = 1, \cdots, k$, we have a polynomial expansion

$$e^{(A - \lambda_{r_j} I)t} v_{r_j} = \sum_{h_j=0}^{m(j)-1} (A - \lambda_{r_j} I)^{h_j} t^{h_j} v_{r_j}, \quad (3.20)$$

where $m(j)$ is the nilpotency index of v_{r_j}.

Expanding each v_{r_j} in terms of the standard basis $\{x_1, \cdots, x_n\} \subset \mathbb{K}^n$ we get an expansion

$$e^{(\hat{A} - \mu I)t} w_\mu = \sum_{1 \le i_1 < \cdots < i_k \le n} f_{(i_1, \cdots, i_k)}(t) x_{i_1} \wedge \cdots \wedge x_{i_k}. \quad (3.21)$$

In this summation $f_{(i_1, \cdots, i_k)}(t)$ are the Plücker coordinates of the vector $e^{(\hat{A} - \mu I)t} w_\mu$ and we will abbreviate them by $f_{\underline{i}}(t)$. By (3.20) it follows that $f_{\underline{i}}(t)$ are all polynomials of degree at most $\sum_{j=1}^k m(j) - k$.

If fact we can say more. Differentiating both sides in (3.19) $m - 1$ times and substituting $t = 0$ results in the eigenvector $(A - \mu I)^{m-1} w_\mu$ on the left side of the equality sign. On the right side this operation results in

$$(m-1)! \sum_{1 \le i_1 < \cdots < i_k \le n} g_{\underline{i}} x_{i_1} \wedge \cdots \wedge x_{i_k},$$

where $g_{\underline{i}}$ is the coefficient of the monom t^{m-1} in the polynomial $f_{\underline{i}}(t)$. By definition we have $(A - \mu I)^{m-1} w_\mu \neq 0$ and $(A - \mu I)^m w_\mu = 0$. We conclude that each polynomial $f_{\underline{i}}(t)$ has degree at most $m - 1$ and some coefficients $g_{\underline{i}}$ are nonzero. In addition note that the vector $e^{(\hat{A} - \mu I)t} w_\mu$ is a decomposable vector at all time $t \ge 0$ and the Plücker relations (compare to [5, Section 3])

$$\text{(QR)} \quad \sum_{\kappa=1}^{p+1} (-1)^\kappa \cdot f_{(i_1, \cdots, i_{p-1}, j_\kappa)}(t) \cdot f_{(j_1, \cdots, \hat{j}_\kappa, \cdots, j_{p+1})}(t) = 0 \quad (3.22)$$

have therefore to be satisfied for all $t \ge 0$ as well. By doing the same argument as in [5, Theorem 3.6 and Example 3.7] we conclude that the Plücker coordinates are also satisfied for the coordinates g_i. But this means that $(A-\mu I)^{m-1}w_\mu$ is a decomposable eigenvector which we denote by

$$\beta_1 \wedge \cdots \wedge \beta_k.$$

Consider once more the eigenspace decomposition

$$\alpha_1 \wedge \cdots \wedge \alpha_k = \sum_\lambda w_\lambda$$

as induced by the decomposition (3.15). Let $m(\lambda)$ be the nilpotency index of w_λ and define the operator

$$f(\hat{A}) := (A-\mu I)^{m-1} \prod_{\lambda \neq \mu} (A-\lambda I)^{m(\lambda)}. \qquad (3.23)$$

A direct calculation shows that

$$f(\hat{A})\alpha_1 \wedge \cdots \wedge \alpha_k = \prod_{\lambda\neq\mu}(\mu-\lambda)^{m(\lambda)}\beta_1 \wedge \cdots \wedge \beta_k. \qquad (3.24)$$

We conclude that the decomposable eigenvector $\beta_1 \wedge \cdots \wedge \beta_k$ is in the unobservable subspace U of the system (3.3) and this implies 2).

Actually we have shown more: $\beta_j \in V_{r_j}$ and if the coefficient r_j is repeated m times in the set $\{r_1, \cdots, r_k\}$ then β_j has nilpotency index at most m. But this means that

$$(A-\lambda_{r_1} I)\cdots(A-\lambda_{r_k} I)\beta_j = 0$$

for $j = 1, \cdots, k$.

By linear dependence of the set $\{C\beta_1, \cdots, C\beta_k\}$ there exist scalars $c_1, \cdots, c_k$ not all zero such that

$$c_1 C\beta_1 + \cdots + c_k C\beta_k = 0.$$

But then

$$w \triangleq c_1\beta_1 + \cdots + c_k\beta_k$$

has all required properties for 1). Q.E.D.

A direct consequence is the following Lemma whose proof is clear.

Lemma 13: If for any set of eigenvalues $\lambda_1, \cdots, \lambda_k$ of A one has

$$\text{rank}\begin{bmatrix}(A-\lambda_1 I)\cdots(A-\lambda_k I)\\ C\end{bmatrix} = n \qquad (3.25)$$

then there is no real decomposable vector in the unobservable subspace U of the system (3.3).

Remark 14: In general the converse is not true as it is demonstrated through an example in [6].

IV. Proof of the Main Theorem

Proof: We first show the sufficiency of the criterion (2.1). Let $P_0, Q_0 \subseteq \mathbb{K}^n$ be two d-dimensional planes with $P_0 \neq Q_0$. Let $q_0 \in Q_0$ be a point having the property that $q_0 \notin P_0$. Let $\{x_0, \cdots, x_d\} \subseteq P_0$ be a set of points chosen in such a way that the decomposable vector

$$w \triangleq (q_0 - x_0) \wedge (x_1 - x_0) \wedge \cdots \wedge (x_d - x_0)$$

is nonzero. If the rank condition (2.1) holds, it follows from Proposition 12 and Lemma 13 that there is no decomposable vector in the unobservable subspace U of "the augmented system" (3.3). It therefore follows that

$$\hat{C}w(t) = (Ce^{At}q_0 - Ce^{At}x_0) \wedge (Ce^{At}x_1 - Ce^{At}x_0) \cdots \wedge (Ce^{At}x_d - Ce^{At}x_0)$$

is nonzero for all time t with the exception of a set of measure zero. But then we have that $Ce^{At}q_0 \notin Ce^{At}P_0$ for almost all time t. In other words $Ce^{At}Q_0 \neq Ce^{At}P_0$.

In order to prove the necessity part assume that all eigenvalues of A are in $\mathbb{K}$. Assume that there is a set of eigenvalues $\lambda_0, \cdots, \lambda_d$ of A such that the rank condition (2.1) is not satisfied. Furthermore assume that for any set of eigenvalues $\mu_0, \cdots, \mu_{d-1}$ of A

$$\text{rank}\begin{bmatrix}(A-\mu_0 I)\cdots(A-\mu_{d-1} I)\\ C\end{bmatrix} = n. \qquad (4.1)$$

If this (technical) condition is not satisfied we will be able to show at the end of the proof that $(d-1)$-dimensional subspaces cannot be observed in general.

By Proposition 12 there exists a nonzero, decomposable vector $x_0 \wedge x_1 \wedge \cdots \wedge x_d$ in the unobservable subspace U of "the augmented system" (3.3) (assuming $k = d+1$). Define $V := \text{span}\{x_0, \cdots, x_d\}$ and let $P_0, Q_0 \subset V$ be two d-dimensional subspaces satisfying $P_0 \neq Q_0$. By the assumption it follows that

$$\text{span}\{Ce^{At}x_0, \cdots, Ce^{At}x_d\}$$

is a d-dimensional subspace for t almost everywhere. But this means that the two different subspaces P_0 and Q_0 in $\mathbb{K}^n$ produce the same moving plane $Ce^{At}P_0 = Ce^{At}Q_0$ in $\mathbb{K}^p$ for all time t with the possible exception of a set of measure zero.

Assume now that (4.1) is not satisfied and let $\tilde{d}$ be the largest integer having the property that

$$\text{rank}\begin{bmatrix}(A-\lambda_0 I)\cdots(A-\lambda_{\tilde{d}} I)\\ C\end{bmatrix} < n \qquad (4.2)$$

for some eigenvalues $\lambda_0, \cdots, \lambda_{\tilde{d}}$ but

$$\text{rank}\begin{bmatrix}(A-\mu_0 I)\cdots(A-\mu_{\tilde{d}-1} I)\\ C\end{bmatrix} = n \qquad (4.3)$$

for all eigenvalues $\mu_0, \cdots, \mu_{\tilde{d}-1}$ of A. Using the same argument as before one shows the existence of two subspaces $\tilde{P}_0$ and $\tilde{Q}_0$ of dimension $\tilde{d}$ which cannot be distinguished in the observation. This completes the proof. Q.E.D.

Acknowledgment

The first author would like to thank Prof. C. Martin for many interesting advice on the perspective observability problem. Comments made by three anonymous referees on an earlier version of this paper is gratefully acknowledged. Finally we would like to thank A. A. Stoorvogel for making numerous suggestions and comments to improve the presentation of the paper.

References

[1] B. K. Ghosh, M. Jankovic, and Y. T. Wu, "Perspective problems in system theory and its application to machine vision," *J. Math. Syst., Estimation Contr.*, vol. 4, no. 1, pp. 3–38, 1994.

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics.* New York: Springer-Verlag, 1978.

[3] M. L. J. Hautus, "Controllability and observability condition of linear autonomous systems," *Ned. Akad. Wetenschappen, Proc. Ser. A*, vol. 72, pp. 443–448, 1969.

[4] B. K. P. Horn, *Robot Vision.* Cambridge, MA: MIT Press, 1986.

[5] J. Rosenthal, "On dynamic feedback compensation and compactification of systems," *SIAM J. Contr. Optim.*, vol. 32, no. 1, pp. 279–296, 1994.

[6] W. P. Dayawansa, B. K. Ghosh, C. Martin, and X. Wang, "A necessary and sufficient condition for the perspective observability problem," *Syst. Contr. Lett.*, to be published.