University of Groningen

Bachelor Project

Applied Mathematics

Geometric Characterisation of Passive Systems

Author:
Gniewko Neyman
(S4915313)

1st Supervisor:
M. K. Camlibel
2nd Supervisor:
B. Besselink

July 3, 2024

rijksuniversiteit groningen

Abstract

This thesis applies the geometric approach to control theory of linear systems to the class of passive linear systems. The assumption of passivity of a given linear system proves to be so restrictive, that explicit characterisations of the weakly unobservable subspace, the strongly reachable subspace, as well as their intersection and sum, can be given in terms of subspaces of the matrices A, B, C, D that fully encompass the dynamics of the system. The thesis provides these, as well as their derivations, and the relevant restrictions that passivity imposes. Lastly, these results are applied to the nine-fold canonical decomposition of a general linear system, from which a six-fold decomposition is derived.

Contents

1 Introduction

The study of passivity in systems theory originated from modelling of physical systems, particularly electrical circuits. Governed by physical laws, such systems often have associated quantities that are internally conserved, and such behavior of conservation was sought to be exploited in order to derive various properties. Such notions were generalised into the concept of a theoretical storage function as a non-negative function of the state of the system, whose change in time is bounded from above by a given supply rate. The study of passivity then, became the study of such systems when defining the supply rate to be the inner product between the input and output of the system. The application of passivity to linear time-invariant systems, done by Jan C. Willems, demonstrated that all storage functions of passive linear systems can be bounded from above and below by quadratic storage functions. As such the study of passivity of linear systems is equivalent to the study of quadratic storage functions associated with such systems.

Separately, much work has been done on the study of general linear systems. Of particular importance to the thesis is the geometric approach to linear systems, whose development was motivated by the disturbance decoupling problem. Geometric properties, compiled in "Control theory for Linear Systems" by H. Trentelman, A. Stoorvogel and M. Hautus, prove to be instrumental in providing properties and solutions to problems by merely considering subspaces associated with the matrices describing the linear system.

Insofar as to the author's knowledge, a geometric approach to passive linear systems was never applied, as such the main objective of this thesis is to do so. Utilising basic properties that passivity imposes on linear systems, the thesis provides an explicit characterisation of the strongly reachable subspace ($\mathcal{S}^*$), the weakly unobservable subspace ($\mathcal{V}^*$), their intersection ($\mathcal{R}^*$), and their sum ($\mathcal{N}^*$) in terms of subspaces associated with the matrices (A, B, C, D) governing all behavior of the linear system. The thesis demonstrates that

$$\mathcal{S}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D \tag{1}$$

$$\mathcal{V}^* = C^{-1} \operatorname{im} D \tag{2}$$

$$\mathcal{R}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle \tag{3}$$

$$\mathcal{N}^* = \mathbb{R}^n. \tag{4}$$

Lastly, the thesis concludes by applying the derived subspace representations to the nine-fold canonical decomposition of a general linear system, attained by H. Aling and J. M. Schumacher. Passivity then reduces such a decomposition to a six-fold one.

2 Preliminaries

2.1 Linear Algebra

Perhaps unsurprisingly, the primary tool used to study linear systems turns out to be linear algebra, and this is no different in this thesis. We will however restrict our scope to vector spaces over the real numbers, and denote such spaces as $\mathbb{R}^n$ where the exponent n signifies the dimension of the space. Similarly we shall denote matrix spaces over the reals by $\mathbb{R}^{n\times m}$ where m and n denote the number of rows and columns respectively. The transpose of M will be denoted with the superscript M^T.

The thesis assumes a thorough understanding of linear algebra and linear systems, as well as a grasp of the most important results from calculus. While calculus is employed quite sparingly, linear algebra dominates the thesis, and so we begin by introducing relevant definitions and results from the field.

2.1.1 Subspaces of a Matrix

For the following definitions of this subsection let $M \in \mathbb{R}^{n\times m}$.

Definition 2.1 (Image)**.** *The image of M, denoted* $\operatorname{im} M$*, is the set of all vectors $y \in \mathbb{R}^n$, such that,*

$$y = Mx, \text{ for some } x \in \mathbb{R}^m.$$

Definition 2.2 (Kernel)**.** *The kernel of M, denoted* $\ker M$*, is the set of all vectors $x \in \mathbb{R}^m$ such that,*

$$Mx = 0.$$

The image and kernel of a matrix along with the image and kernel of its transpose are commonly referred to as the four fundamental subspaces associated with a matrix. They encompass the most important behaviors of the matrix as a map from one vector space to another, and are the building blocks upon which later subspace characterisation is attained.

The introduction of the image and kernel of the transpose as fundamental might at first seem arbitrary, but it turns out they are intricately linked through the concept of the orthogonal complement.

Definition 2.3 (Orthogonal complement)**.** *Let $\mathcal{V} \subseteq \mathbb{R}^n$ be a subspace. The orthogonal complement, denoted with the superscript $\mathcal{V}^\perp$ is defined as the set of all y such that*

$$x^T y = 0, \text{ for all } x \in \mathcal{V}.$$

Lemma 2.4. *The kernel of a matrix M is equal to the orthogonal complement of the image of the transpose of M, i.e.* $\ker M = (\operatorname{im} M^T)^\perp$.

Proof. Let x denote an arbitrary vector in $\ker M$, and $M^T y$ an arbitrary vector of the image of M^T. The inner product between them, then is given as

$$x^T M^T y$$

which we can rearrange as

$$(Mx)^T y$$

which when considering that x is a member of the kernel of M yields

$$x^T M^T y = 0^T y = 0$$

which shows that

$$\ker M \subseteq (\operatorname{im} M^T)^\perp.$$

To prove the lemma we must also show the reverse inclusion. As such consider an arbitrary element x in the orthogonal complement of $\operatorname{im} M^T$. We have $x^T M^T y = 0$ which we again rearrange as

$$(Mx)^T y = 0.$$

Now, the above statement holds for all possible vectors y which would imply $Mx = 0$, and so x is a member of the kernel of M, as such

$$(\operatorname{im} M^T)^\perp \subseteq \ker M.$$

Which proves that

$$\ker M = (\operatorname{im} M^T)^\perp.$$

Apart from the fundamental subspaces, we would like to capture the behavior of linear maps when we restrict our domain or co-domain. To this extent we define the image and inverse image of a subspace in relation to a matrix M.

Definition 2.5 (Image of a subspace)**.** *Let $\mathcal{V} \subseteq \mathbb{R}^m$. The subspace $M\mathcal{V} \subseteq \mathbb{R}^n$ is defined as the set of all vectors y such that,*

$$Mx = y, \text{ for some } x \in \mathcal{V}.$$

Definition 2.6 (Inverse image of a subspace)**.** *Let $\mathcal{W} \subseteq \mathbb{R}^n$. The inverse image is denoted as the subspace $M^{-1}\mathcal{W} \subseteq \mathbb{R}^m$ and is the set of all vectors x such that,*

$$Mx = y, \text{ for some } y \in \mathcal{W}$$

Note here that M need not be invertible despite the notation used. It does however allow us to capture the essence of invertibility. For example if we have given that $Mx = y$, then the inverse image $M^{-1}\{y\} = \{x\} + \ker M$ which proves to be a useful property when characterising more complicated subspaces.

It turns out that, as with our four fundamental subspaces, the image and inverse image of a subspace are also related to each other by the orthogonal complement.

Lemma 2.7. *The orthogonal complement of the image of a subspace $\mathcal{V}$ under multiplication by M is equal to the inverse image of its orthogonal complement under the transpose of M, i.e. $(M\mathcal{V})^\perp = (M^T)^{-1}\mathcal{V}^\perp$.*

Proof. Consider an element x in $(M\mathcal{V})^\perp$, then by definition

$$\begin{aligned}
x^T M v &= 0, \text{ for all } v \in \mathcal{V} \\
(M^T x)^T v &= 0, \text{ for all } v \in \mathcal{V} \\
M^T x &\in \mathcal{V}^\perp \\
x &\in (M^T)^{-1}\mathcal{V}^\perp \\
(M\mathcal{V})^\perp &\subseteq (M^T)^{-1}\mathcal{V}^\perp.
\end{aligned}$$

Now to demonstrate the reverse inclusion, consider x to be an element of $(M^T)^{-1}\mathcal{V}^\perp$, then by definition, we have:

$$\begin{aligned}
M^T x &\in \mathcal{V}^\perp \\
(M^T x)^T v &= 0, \text{ for all } v \in \mathcal{V} \\
x^T (Mv) &= 0, \text{ for all } v \in \mathcal{V} \\
x &\in (M\mathcal{V})^\perp \\
(M^T)^{-1}\mathcal{V}^\perp &\subseteq (M\mathcal{V})^\perp.
\end{aligned}$$

Combining both derived inclusions yields equality, proving that

$$(M\mathcal{V})^\perp = (M^T)^{-1}\mathcal{V}^\perp.$$

While the so far covered definitions and lemmas are results from general linear algebra, we'd like to touch upon concepts much more interconnected with linear systems. To this end we define subspace invariance under matrix multiplication.

Definition 2.8 (Subspace invariance)**.** *Let M be a square matrix, ($M \in \mathbb{R}^{n\times n}$). A subspace $\mathcal{V}$ is called M-invariant if*

$$M\mathcal{V} \subseteq \mathcal{V}.$$

The property of invariance shows up quite frequently when dealing with problems in linear systems. However, for arbitrarily chosen matrices and subspaces, invariance is not the norm. As such, we'd like to encapuslate this property of invariance for more general subspaces that need not themselves be invariant.

Definition 2.9. *The smallest M-invariant subspace containing $\mathcal{V}$ is denoted by $\langle M \mid \mathcal{V}\rangle$ and is the unique subspace satisfying the following three properties:*

1. $\mathcal{V} \subseteq \langle M \mid \mathcal{V}\rangle$
2. $M\langle M \mid \mathcal{V}\rangle \subseteq \langle M \mid \mathcal{V}\rangle$
3. *If $\mathcal{N}$ is a M-invariant subspace containing $\mathcal{V}$, $\langle M \mid \mathcal{V}\rangle \subseteq \mathcal{N}$*

Definition 2.10. *The largest M-invariant subspace contained in $\mathcal{V}$ is denoted by $\langle \mathcal{V} \mid M\rangle$ and is the unique subspace satisfying the following three properties:*

1. $\langle \mathcal{V} \mid M\rangle \subseteq \mathcal{V}$
2. $M\langle \mathcal{V} \mid M\rangle \subseteq \langle M \mid \mathcal{V}\rangle$
3. *If $\mathcal{N}$ is a M-invariant subspace contained in $\mathcal{V}$, $\mathcal{N} \subseteq \langle \mathcal{V} \mid M\rangle$*

Invariance is perhaps the quintessential tool in the study of linear systems, and we will see it appear throughout the thesis.

2.1.2 Definite Matrices

The concept of matrix definiteness in the context of linear systems arose from the physical realities that such systems often model. In this section, we introduce the most important properties of such matrices.

Definition 2.11 (Symmetric matrix)**.** *A square matrix is called symmetric if it is equal to its transpose, i.e.* $M = M^T$*. We denote the set of all symmetric matrices of dimension* n *as* $\mathbb{S}^n$

Definition 2.12 (Definite matrix)**.** *Let* M *be in* $\mathbb{S}^n$*.* M *is called positive semidefinite if*

$$x^T M x \geqslant 0, \text{ for all } x \in \mathbb{R}^n.$$

It is called positive definite if instead this inequality can be made strict, i.e.

$$x^T M x > 0, \text{ for all } x \in \mathbb{R}^n \text{ and } x \neq 0.$$

The set of all positive semidefinite matrices is denoted as $\mathbb{S}^n_+$*, and of all positive definite matrices as* $\mathbb{S}^n_{++}$*. A matrix* M *is called negative (semi)definite if* $-M$ *is positive (semi)definite.*

The function $f(x) = x^T M x$ is called the quadratic form associated with M. The fact that it is always non negative for positive (semi)definite matrices M is sought after as it allows for the modelling of non-negative physical quantities, such as energy.

Such matrices posses quite useful properties.

Lemma 2.13. *Let* M *be in* S^n_+*. Then* $x \in \ker M$ *if and only if* $x^T M x = 0$*.*

Proof. First assume $x \in \ker M$. This implies $Mx = 0$, which implies $x^T M x = 0$.

Now assume that $x^T M x = 0$, since M is positive semidefinite, $M = N^T N$ for some matrix N. Plugging this into our assumption yields

$$\begin{aligned} 0 &= x^T N^T N x \\ &= ||Nx||. \end{aligned}$$

This in turn, necessitates $Nx = 0$, which when multiplied from the left by N^T grants us our desired result since $Mx = N^T N x = 0$ and so x is in the kernel of M.

The above effectively shows the equivalence between a vector annihilating the quadratic form associated with a matrix and that same vector annihilating the matrix itself. This, in turn, allows us to show that a vector belongs to the kernel of a matrix by instead showing the quadratic form is equal to zero, which is often much easier.

Additionally, throughout the thesis, positive semidefinite matrices will frequently obtain the form $M+M^T$, which gives rise to even more properties.

Lemma 2.14. *Let M be such that $M + M^T \in S^n_+$. Then,*

$$\begin{aligned} &i.\ \ker M \subseteq \ker(M + M^T) \\ &ii.\ \ker M = \ker M^T \end{aligned}$$

Proof. We begin with a proof of (i). Take an arbitrary element $x \in \ker M$, then $Mx = 0$. Now note that

$$\begin{aligned} x^T(M + M^T)x &= x^T Mx + x^T M^T x \\ &= x^T Mx + x^T M^T x \\ &= x^T(Mx) + (Mx)^T x \\ &= 0. \end{aligned}$$

As such $x^T(M + M^T)x = 0$. Now since $M + M^T \in S^n_+$, we use Lemma 2.13, which implies $x \in \ker(M + M^T)$

For the proof of (ii). Note that by (i), we have $x \in \ker M \Rightarrow x \in \ker(M + M^T)$, so $(M + M^T)x = 0$. Expanding gives $Mx + M^T x = 0$ which simplifies to $M^T x = 0$ and so $x \in \ker M^T$. The reverse inclusion also holds by symmetry of the argument with respect to transposition. Therefore we conclude that

$$\ker M = \ker M^T.$$

From the above, we can see that such a construction imposes symmetry of the kernel with respect to transposition.

2.2 Linear Input-State-Output Systems

Having introduced the relevant concepts from linear algebra, we turn our attention to the main mathematical object of study in this thesis, namely the linear system.

Definition 2.15 (Linear Input-State-Output Systems)**.** *A linear input-state-output system, denoted by $\Sigma(A, B, C, D)$, is a set of first-order differential equations with respect to continuous time ($t \in \mathbb{R}$) of the form:*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

Here $x(t), u(t), y(t)$ are vector-valued functions in time where $x(t) \in \mathbb{R}^n$ is called the state, $u(t) \in \mathbb{R}^m$ is called the input, $y(t) \in \mathbb{R}^p$ is called the output, and $\dot{x}(t) := \frac{d}{dt}x(t)$, and A, B, C, D are matrices of appropriate dimensions. For a given initial condition $x(0) = x_0$ and locally integrable input function $u(t)$, the solution $x_{x_0,u}(t)$ is called the state trajectory, with the corresponding solution $y_{x_0,u}(t)$ being called the output trajectory.
(NB. Once defined, often $\Sigma(A, B, C, D)$ will be denoted simply as Σ, and $x(t), u(t), y(t)$ as x, u, y)

What is remarkable, is that while the definition of a linear system is reliant on the notion of the derivative, it is often the case that we can describe various properties of such a system solely in terms of linear algebraic properties. This is commonly referred to as the geometric approach, the core concepts of which we introduce now.

2.2.1 Fundamental Subspaces of $\Sigma(A, B, C, D)$

In the following section, we will list the definitions used in the geometric approach to the study of linear systems, adapted from Trentelman, Stoorvogel, and Hautus 2002.

Let $\Sigma(A, B, C, D)$ be a system as in Definition 2.15.

Definition 2.16 (Reachable subspace)**.** *The subspace $\langle A \mid \operatorname{im} B\rangle$ is called the reachable subspace of Σ. It is the smallest A-invariant subspace containing $\operatorname{im} B$ and has the property that, $x \in \langle A \mid \operatorname{im} B\rangle$ if and only if there exists an input function $u(t)$ and time $t_1 \geqslant 0$, such that $x_{0,u}(t_1) = x$.*

Definition 2.17 (Unobservable subspace)**.** *The subspace $\langle \ker C \mid A\rangle$ is called the unobservable subspace of Σ. It is the largest A-invariant subspace contained in $\ker C$ and has the property that, $x_0 \in \langle \ker C \mid A\rangle$ if and only if $y_{x_0,0}(t) = 0$ for all times t.*

Definition 2.18 (Weakly unobservable subspace)**.** *The weakly unobservable subspace of Σ, denoted by $\mathcal{V}^*$, is the largest subspace $\mathcal{V}$ of $\mathbb{R}^n$ such that:*

$$\begin{bmatrix} A \\ C \end{bmatrix} \mathcal{V} \subseteq (\mathcal{V} \times 0) + \operatorname{im} \begin{pmatrix} B \\ D \end{pmatrix}.$$

The weakly unobservable subspace has the property that for all $x_0 \in \mathcal{V}^*$, there exists an input function $u(t)$ such that $y_{x_0,u}(t) = 0$ for all times $t \geqslant 0$.

Definition 2.19 (Strongly reachable subspace)**.** *The strongly reachable subspace of Σ, denoted by $\mathcal{S}^*$, is the smallest subspace $\mathcal{S}$ of $\mathbb{R}^n$ such that:*

$$\begin{bmatrix} A & B \end{bmatrix} \left[(\mathcal{S} \times \mathbb{R}^p) \cap \ker \begin{pmatrix} C & D \end{pmatrix} \right] \subseteq \mathcal{S}.$$

The strongly reachable subspace can be seen as the equivalent of the reachable subspace when restricting inputs to the set of impulse functions (see Trentelman, Stoorvogel, and Hautus (2002) p. 368).

Proposition 2.20. *(Duality between the weakly unobservable subspace and the strongly reachable subspace)*
We define the dual of a linear system $\Sigma(A, B, C, D)$ as $\Sigma^T := \Sigma(A^T, C^T, B^T, D^T)$, then there is a strong relation between the subspaces of the dual with its original, namely:

$$\mathcal{S}^*(\Sigma)^\perp = \mathcal{V}^*(\Sigma^T)$$

and

$$\mathcal{V}^*(\Sigma)^\perp = \mathcal{S}^*(\Sigma^T).$$

See Trentelman, Stoorvogel, and Hautus 2002 p.186 for details.

Definition 2.21 (The controllable weakly unobservable subspace)**.** *The controllable weakly unobservable subspace, denoted by $\mathcal{R}^*$, is the intersection of the weakly unobservable subspace and the strongly reachable subspace.*

$$\mathcal{R}^* := \mathcal{V}^* \cap \mathcal{S}^*.$$

Definition 2.22 (The distributionally weakly unobservable subspace)**.** *The distributionally weakly unobservable subspace, denoted by $\mathcal{W}^*$,is the sum of the weakly unobservable subspace with the strongly reachable subspace.*

$$\mathcal{W}^* := \mathcal{V}^* + \mathcal{S}^*.$$

These properties allow for a much easier analysis of linear systems. They prove the existence of specific input functions driving desired system behavior without the need to explicitly define them. If we treat the system Σ as a map from the space of input functions to the space of output functions, these subspaces allow us to prove injectivity and surjectivity of such a map.

2.2.2 Passivity

Having defined the linear system, and concepts associated with it, we proceed to the restriction that we are interested on imposing onto the system, that being passivity. Passivity arises quite intuitively when modelling physical systems, where we deal with certain conserved quantities, most often energy. These laws of conservation are often only present implicitly in such models, and so we'd like to quantify such behavior much more concretely.

Definition 2.23 (Passive system)**.** *A system Σ, whose input and output spaces are of equal dimenstion, is called **passive** if there exists a $P \geqslant 0$ such that the **dissipativity inequality** (DI)*

$$\frac{1}{2}x^T(t_0)Px(t_0) + \int_{t_0}^{t_1} u^T(\tau)y(\tau)d\tau \geqslant \frac{1}{2}x^T(t_1)Px(t_1),$$

holds for all times t_0, t_1 with $t_0 \leqslant t_1$, and for all trajectories (u, x, y) of Σ.

We call $w(t) := u^T(t)y(t)$ the supply rate function, and $s(x) := \frac{1}{2}x^T Px$ the associated storage function of the system. Passivity then tells us that such a storage function is internally conserved, and only increasing in the presence of an external supply rate.

While the definitions for the supply rate and storage might seem arbitrary, it turns out that, specifically for passive linear electric circuits, one can always derive an associated linear system model such that the above definition of passivity holds.

Such a definition, however, is often impractical to deal with and so we seek out an easier to work with characterisation.

Proposition 2.24 (Linear matrix inequality of passivity)**.** *A linear system $\Sigma(A, B, C, D)$ is passive if and only if there exists $P \geqslant 0$ such that:*

$$\begin{bmatrix} 0 & C^T \\ C & D + D^T \end{bmatrix} - \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \geqslant 0.$$

Proof. We first assume the system is passive, as such we have:

$$\frac{1}{2}x^T(t_0)Px(t_0) + \int_{t_0}^{t_1} u^T(\tau)y(\tau)d\tau \geqslant \frac{1}{2}x^T(t_1)Px(t_1),$$

for all times t_0, t_1 such that $t_0 \leqslant t_1$, and for all trajectories (u, x, y) of Σ.

For simplicity we will omit the notations of $x(t), u(t), y(t)$ as functions of t, instead denoting them simply as x, u, y, except when evaluated at a specific point in time. We rearrange to obtain:

$$\int_{t_0}^{t_1} u^T(\tau)y(\tau)d\tau \geqslant \frac{1}{2}x^T(t_1)Px(t_1) - \frac{1}{2}x^T(t_0)Px(t_0)$$

Now by the fundamental theorem of calculus, for a given function $W(t)$ such that $\frac{d}{dt}W(t) = u^T(t)y(t)$, we can rewrite the above as:

$$W(t_1) - W(t_0) \geqslant \frac{1}{2}x^T(t_1)Px(t_1) - \frac{1}{2}x^T(t_0)Px(t_0)$$

Now take $t_0 \neq t_1$ such that $t_1 - t_0 > 0$, then dividing both sides of the inequality gives:

$$\frac{W(t_1) - W(t_0)}{t_1 - t_0} \geqslant \frac{\frac{1}{2}x^T(t_1)Px(t_1) - \frac{1}{2}x^T(t_0)Px(t_0)}{t_1 - t_0}.$$

We now take the limit of both sides as t_1 approaches t_0 to obtain

$$\lim_{t_1 \to t_0} \frac{W(t_1) - W(t_0)}{t_1 - t_0} \geqslant \lim_{t_1 \to t_0} \frac{\frac{1}{2}x^T(t_1)Px(t_1) - \frac{1}{2}x^T(t_0)Px(t_0)}{t_1 - t_0}$$

which when recalling the definition of the derivative simplifies to

$$\frac{d}{dt}W(t_0) \geqslant \frac{d}{dt}(\frac{1}{2}x^T Px)(t_0).$$

Recalling how we defined $W(t)$ we finally we obtain

$$u^T(t_0)y(t_0) \geqslant \frac{d}{dt}(\frac{1}{2}x^T Px)(t_0).$$

However, the choice of t_0 was arbitrary, so the statement holds for all t, and as such:

$$\begin{aligned} u^T y &\geqslant \frac{d}{dt}(\frac{1}{2}x^T Px) \\ 2u^T y &\geqslant \frac{d}{dt}(x^T Px) \\ u^T y + y^T u &\geqslant \dot{x}^T Px + x^T P\dot{x}. \end{aligned}$$

Now, recall the structure of a linear system, namely:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du. \end{aligned}$$

Plugging in the expressions for $\dot{x}$ and y into the derived inequality yields:

$$u^T(Cx + Du) + (Cx + Du)^T u \geqslant (Ax + Bu)^T Px + x^T P(Ax + Bu).$$

Expanding and rearranging terms gives us

$$0 \geqslant x^T(A^T P + PA)x + u^T(B^T P - C)x + x^T(PB - C^T)u - u^T(D + D^T)u$$

which can be nicely rewritten in block form as:

$$0 \geqslant \begin{bmatrix} x^T & u^T \end{bmatrix} \begin{bmatrix} A^T P + PA & PB - C^T \\ B^T P - C & -(D + D^T) \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

or equivalently as:

$$\begin{bmatrix} x^T & u^T \end{bmatrix} \left(\begin{bmatrix} 0 & C^T \\ C & D + D^T \end{bmatrix} - \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \right) \begin{bmatrix} x \\ u \end{bmatrix} \geqslant 0.$$

Now note that passivity applies to all possible states and inputs, as such the above statement applies to any arbitrary vector $\begin{bmatrix} x \\ u \end{bmatrix} \in \mathbb{R}^{n+m}$ showing that indeed:

$$\begin{bmatrix} 0 & C^T \\ C & D + D^T \end{bmatrix} - \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \geqslant 0.$$

The "if" statement follows analogously, by following the algebraic steps in inverse order, until one obtains:

$$u^T y \geqslant \frac{d}{dt}(\frac{1}{2} x^T P x).$$

Which when integrating from t_0 to t_1, yields:

$$\begin{aligned} \int_{t_0}^{t_1} u^T y d\tau &\geqslant \int_{t_0}^{t_1} \frac{d}{d\tau}(\frac{1}{2} x^T P x) d\tau \\ &\geqslant \frac{1}{2} x^T(t_1) P x(t_1) - \frac{1}{2} x^T(t_0 P x(t_0) \end{aligned}$$

which when rearranged gives

$$\frac{1}{2} x^T(t_0) P x(t_0) + \int_{t_0}^{t} u^T y d\tau \geqslant \frac{1}{2} x^T(t_1) P x(t_1)$$

as desired.

This theorem transports the definition of passivity from one that relies on calculus to one that encapsulates passivity with linear algebra. This in turn allows for subspace characterisation in a much more simple manner.

However, due to the sheer size of the block matrix associated with the passivity inequality, it would be impractical to work with. As such, we introduce simplified notation.

Definition 2.25 (Output Matrix and State map). *Let $\Sigma(A, B, C, D)$ be a passive system. We define the output matrix S as*

$$S := \begin{bmatrix} 0 & C^T \\ C & D + D^T \end{bmatrix}$$

and the state map $L : \mathbb{S}^n \to \mathbb{S}^{n+m}$ *as*

$$L(P) := \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix}$$

from which the dissipativity inequality for linear systems can be simply rewritten as

$$S - L(P) \geqslant 0.$$

For a given Σ *the set of all matrices* P *satisfying the dissipativity inequality is denoted as* $\mathcal{P} \subseteq \mathbb{S}^n_+$, *more concretely*

$$\mathcal{P} := \{P \geqslant 0 : S - L(P) \geqslant 0\}.$$

Having obtained a linear algebraic formulation of passivity, we now turn our attention to the various consequences that arise upon its assumption. Indeed, the assumption of passivity imposes many restrictions on the matrices (A, B, C, D) of a linear system.

Proposition 2.26 (Consequences of passivity)**.** *Let* $\Sigma(A, B, C, D)$ *be a passive system. The following statements hold:*

(i) $D + D^T \geqslant 0$

(ii) $\ker D \subseteq \ker(D + D^T)$

(iii) $\ker D = \ker D^T$

(iv) $\operatorname{im} D = \operatorname{im} D^T$.

Furthermore, for $P \in \mathcal{P}$

(v) $(PB - C^T)\ker(D + D^T) = \{0\}$

(vi) $\ker P \subseteq \ker C$

(vii) $\ker P$ *is* A*-invariant*

(viii) $\ker P \subseteq \langle \ker C \mid A \rangle$.

Proof. (i): Consider $x = \begin{bmatrix} 0 \\ \mu \end{bmatrix}$ for an arbitrary $\mu \in \mathbb{R}^m$, then

$$x^T(S - L(P))x = \mu^T(D + D^T)\mu \geqslant 0$$

which by arbitrary choice of μ implies $(D + D^T) \geqslant 0$.

(ii) and (iii): These statements directly follow from (i) and Lemma 2.14.

(iv): This is a direct consequence of taking the orthogonal complement of (iii) and applying Lemma 2.4.

(v): Consider again $x = \begin{bmatrix} 0 \\ \mu \end{bmatrix}$ this time with $\mu \in \ker(D + D^T)$. Then

$$x^T(S - L(P))x = \mu^T(D + D^T)\mu = 0$$

since by assumption $\mu \in \ker(D + D^T)$. However, since $S - L(P) \geqslant 0$ and $x^T(S - L(P))x = 0$, by Lemma 2.13, $x \in \ker(S - L(P))$ which then implies $(S - L(P))x = 0$. Now

$$\begin{aligned} 0 &= (S - L(P))x \\ &= C^T\mu - PB\mu + (D + D^T)\mu \\ &= (C^T - PB)\mu. \end{aligned}$$

Since $\mu \in \ker(D + D^T)$ was selected arbitrarily, this shows that indeed (v) holds.

(vi): Consider $x = \begin{bmatrix} \xi \\ 0 \end{bmatrix}$ for an arbitrary $\xi \in \ker P$. By algebraic manipulation along with P being symmetric, one can again show that $x^T(S - L(P))x = 0$, which by Lemma 2.13 implies

$$\begin{aligned} 0 &= (S - L(P))x \\ &= (B^T P - C)\xi \\ &= C\xi \end{aligned}$$

and so

$$\xi \in \ker C.$$

(vii): This statement follows by first noting that $A^T P + PA$ is negative semidefinite (Analogous to the proof of (i), along with the fact that for $\xi \in \ker P$, $\xi^T(A^T P + PA)\xi = 0$, as such by lemma 2.13,

$$(A^T P + PA)\xi = 0$$

$$PA\xi = 0$$

$$A\xi \in \ker P$$
$$A \ker P \subseteq \ker P.$$

(viii): This statement directly follows from (vi) and (vii) since $\langle A \mid \operatorname{im} B\rangle$ is the largest A-invariant subspace contained in $\ker C$.

Passivity imposes symmetry on the four fundamental subspaces of D, particularly statement (iii) which implies that if $D\mu = 0$ then $\mu^T D = 0$. Perhaps most importantly however is statement (v) which says that if $D\mu = 0$ then $PB\mu = C^T\mu$. These two statements appear most often throughout the proofs in the following sections, so it is useful to keep them in mind.

We wish to say a bit more about two particular intersections, $B \ker D \cap C^{-1} \operatorname{im} D$ and $C^T \ker D^T \cap (B^T)^{-1} \operatorname{im} D^T$.

Proposition 2.27. *Let $\Sigma(A, B, C, D)$ be a passive system. Then, we have:*

(i) $B \ker D \cap C^{-1} \operatorname{im} D \subseteq \ker P$

(ii) $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle \subseteq \ker P$

(iii) $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle \subseteq \ker C$

(iv) $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle \subseteq C^{-1} \operatorname{im} D$

(v) $(\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle + B \ker D) \cap C^{-1} \operatorname{im} D = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$

Proof. (i): Consider an element $x \in B \ker D \cap C^{-1} \operatorname{im} D$. As it is in the intersection, it obeys the properties of both subspaces, as such we can rewrite x as $x = B\mu$ for some element $\mu \in \ker D$ as well as knowing that there exists some vector u such that $Cx = Du$. Putting both results together yields $CB\mu = Du$, which when multiplying from the left by μ^T, and rearranging, yields $(C^T\mu)^T B\mu = (D^T\mu)^T u$. By statement (iii) of Lemma 2.26 $D^T\mu = 0$, and by statement (v) of Lemma 2.26 $C^T\mu = PB\mu$, which then implies $\mu^T B^T PB\mu = 0$, which when remembering how x was defined yields $x^T Px = 0$, which by Lemma 2.13 shows $x \in \ker P$.

(ii): First note that $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle \subseteq \langle A \mid \ker P\rangle$ by (i). Then by the A-invariance of $\ker P$ (statement (vii) of Lemma 2.26) $\langle A \mid \ker P\rangle \subseteq \ker P$, and so $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$ must also therefore be in $\ker P$.

(iii): The statement follows from (ii) and from noting that $\ker P \subseteq \ker C$ by statement (vi) of Lemma 2.26.

(iv): This statement follows from (iii) and by noting that $\ker C \subseteq C^{-1} \operatorname{im} D$.

(v): While the construction of the statement of (v) might seem overly complicated and arbitrary, this particular subspace appears multiple times in the following section and as such it is worthwhile to characterise it.

The proof of (v) can be easily demonstrated by the closedness of subspaces under addition. Take $x \in (\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D) \cap C^{-1} \operatorname{im} D$, then $x = \eta + \zeta$, for some $\eta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle \subseteq C^{-1} \operatorname{im} D$ and $\zeta \in B \ker D$. However note two key things, firstly x is also in the right subspace of the intersection, so $x \in C^{-1} \operatorname{im} D$, additionally by (iv) $\eta \in C^{-1} \operatorname{im} D$ and so, since subspaces are closed under addition $\zeta = (x - \eta) \subseteq C^{-1} \operatorname{im} D$. Remembering how we defined ζ we just showed that $\zeta \in B \ker D \cap C^{-1} \operatorname{im} D$. Even more remarkable is that this suggests that $\zeta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle$, and so finally $x = \eta + \zeta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle$.

To show the reverse inclusion. take $x \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle$, it is obviously in $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D$ and by (iv) it is also in $C^{-1} \operatorname{im} D$ so it must also be in their intersection, showing inclusion both ways and finalizing the proof.

Proposition 2.28. *Let $\Sigma(A, B, C, D)$ be a passive system. Then*

$$C^T \ker D^T \cap (B^T)^{-1} \operatorname{im} D^T = 0.$$

Proof. Let x be in $C^T \ker D^T \cap (B^T)^{-1} \operatorname{im} D^T$. Then $x = C^T \mu = PB\mu$ for some $\mu \in \ker D^T$ (by statement (iii) and (v) of Lemma 2.26) and there exists a u such that $B^T x = D^T u$. Combining the two equations gives us $B^T PB\mu = D^T u$. Then, multiplying both sides from the left by μ^T and using statement (iii) of Lemma 2.26 yields $\mu^T B^T PB\mu = 0$, However, by Lemma 2.13 this would imply that $PB\mu = 0$. Then x which is equal to $PB\mu$ must also necessarily be zero.

As we will soon see in the following section, these two subspaces alone almost fully explicitly characterise the geometric structure of a passive linear system.

3 Geometric Structure of Passive Systems

In the following section, we consider a passive linear system $\Sigma(A, B, C, D)$ and compute the weakly unobservable, strongly reachable subspaces as well as their sum and intersection.

Theorem 3.1 (Strongly reachable subspace of a passive system)**.** *The strongly reachable subspace of a passive linear system is*

$$\mathcal{S}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D)\rangle + B \ker D.$$

Proof. The proof relies on showing that the above defined subspace obeys the extremal properties of a strongly reachable subspace, namely

$$\begin{bmatrix} A & B \end{bmatrix} \left[(\mathcal{S}^* \times \mathbb{R}^p) \cap \ker \begin{pmatrix} C & D \end{pmatrix} \right] \subseteq \mathcal{S}^*.$$

And it is contained in all other subspaces obeying the above property.

First we prove that the above defined subspace indeed obeys this property. As such take an element s such that

$$s \in \begin{bmatrix} A & B \end{bmatrix} \left[(\mathcal{S}^* \times \mathbb{R}^p) \cap \ker \begin{pmatrix} C & D \end{pmatrix} \right].$$

In other words for some $\eta \in \mathcal{S}^*$ and ζ satisfying $C\eta + D\zeta = 0$ our chosen s takes the form $s = A\eta + B\zeta$. We want to show that $s \in \mathcal{S}^*$ in order for it to satisfy the definition.

First note that, since by assumption $C\eta + D\zeta = 0$, $\eta \in C^{-1} \operatorname{im} D$, so along with our other assumption about η, namely $\eta \in \mathcal{S}^*$, $\eta \in \mathcal{S}^* \cap C^{-1} \operatorname{im} D$. Then by statement (v) of Lemma 2.27 this intersection is equal to $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$. So $\eta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$, but by A-invariance, $A\eta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle \subseteq \mathcal{S}^*$.

After showing that $A\eta \subseteq \mathcal{S}^*$ all that is now left to show is that $B\zeta$ is also in $\mathcal{S}^*$. To do so, recall that $\eta \in \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$. Then by statement (iii) of Lemma 2.27 $\eta \in \ker C$, so $C\eta + D\zeta = 0$ simplifies to $D\zeta = 0$ and as such ζ must be in $\ker D$, and therefore $B\zeta \in B \ker D \subseteq \mathcal{S}^*$.

Putting both results together then, shows that $s = A\eta + B\zeta \in \mathcal{S}^*$.

To finish the proof, we must also show that the $\mathcal{S}^*$ we defined is the smallest such subspace obeying this property. As such assume any other subspace $\mathcal{S}$ satisfies

$$\begin{bmatrix} A & B \end{bmatrix} \left[(\mathcal{S} \times \mathbb{R}^p) \cap \ker \begin{pmatrix} C & D \end{pmatrix} \right] \subseteq \mathcal{S}.$$

Note that by this definition

$$A(S \cap \ker C) \subseteq S.$$

With this fact in hand, let us show that $\mathcal{S}^* \subseteq \mathcal{S}$. Firstly, any subspace contains 0, then

$$\begin{bmatrix} A & B \end{bmatrix} \left[(0 \times \mathbb{R}^p) \cap \ker \begin{pmatrix} C & D \end{pmatrix} \right] = B \ker D \subseteq \mathcal{S}.$$

Now since $B \ker D \cap C^{-1} \operatorname{im} D$ is contained in $B \ker D$ it must also be in $\mathcal{S}$. The key element to the proof is then noting that by statement (iii) of Lemma2.27 $B \ker D \cap C^{-1} \operatorname{im} D \subseteq \ker C$ and so by definition of the strongly reachable subspace, $A(B \ker D \cap C^{-1} \operatorname{im} D) \subseteq \mathcal{S}$. Lastly, by statements (iii) of Lemma 2.27,(vi) and (vii) of Lemma 2.26 $A(B \ker D \cap C^{-1} \operatorname{im} D) \subseteq A \ker P \subseteq \ker P \subseteq \ker C$, and so $A^2(B \ker D \cap C^{-1} \operatorname{im} D) \subseteq \mathcal{S}$.

Repeating this line of argument ad infinitum, we get that:

$$A^k(B \ker D \cap C^{-1} \operatorname{im} D) \subseteq \mathcal{S}, \forall k \in \mathbb{N}$$

$$\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle \subseteq \mathcal{S}.$$

Which when combined with the fact that $B \ker D \subseteq \mathcal{S}$ shows

$$\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D = \mathcal{S}^* \subseteq \mathcal{S}$$

for any other $\mathcal{S}$ satisfying the strongly reachable subspace inclusion.

Theorem 3.2. *The weakly unobservable subspace of a passive linear system is given by*

$$\mathcal{V}^* = C^{-1} \operatorname{im} D.$$

Proof. We exploit duality between the strongly reachable and the weakly unobservable subspace given by the equality of statement 2.20, namely:

$$\mathcal{V}^*(A, B, C, D) = \mathcal{S}^*(A^T, C^T, B^T, D^T)^{\perp}.$$

By Lemmas 2.4 and 2.7 $(C^T \ker D^T)^{\perp} = C^{-1} \operatorname{im} D$ and so the statement $\mathcal{V}^*(A, B, C, D) = C^{-1} \operatorname{im} D$ being true is equivalent to proving

$$\mathcal{S}^*(A^T, C^T, B^T, D^T) = C^T \ker D^T.$$

As such we now prove that $\mathcal{S}^*(A^T, C^T, B^T, D^T) = C^T \ker D^T$. As with our previous proof we must demonstrate that $C^T \ker D^T$ satisfies the subspace inclusion of a strongly reachable subspace and it is contained in all other subspaces satisfying this inclusion.

First we prove that $C^T \ker D^T$ obeys the strongly reachable subspace inclusion, namely:

$$\begin{bmatrix} A^T & C^T \end{bmatrix} \left[(C^T \ker D^T \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \right] \subseteq C^T \ker D^T.$$

Take the vector pair $(\eta, \zeta) \in (C^T \ker D^T \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix}$. By construction $\eta \in C^T \ker D^T$ and $B^T \eta + D^T \zeta = 0$. Combining the two grants $\eta \in C^T \ker D^T \cap (B^T)^{-1} \operatorname{im} D^T$. However by Lemma 2.28 this implies that $\eta = 0$. Then ζ , which by definition satisfies $B^T \eta + D^T \zeta = 0$ must actually satisfy the stricter statement $D^T \zeta = 0$ and so ζ is an element of $\ker D^T$. As such

$$(C^T \ker D^T \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \subseteq 0 \times \ker D^T$$

which after multiplying from the left by the block matrix we indeed get what we sought, i.e.

$$\begin{bmatrix} A^T & C^T \end{bmatrix} \left[(C^T \ker D^T \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \right] \subseteq C^T \ker D^T.$$

Now to show that it is the smallest such subspace, consider any other subspace $\mathcal{S}$ satisfying

$$\begin{bmatrix} A^T & C^T \end{bmatrix} \left[(\mathcal{S} \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \right] \subseteq \mathcal{S}.$$

Since $0 \in \mathcal{S}$,

$$\begin{bmatrix} A^T & C^T \end{bmatrix} \left[(0 \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \right] \subseteq \mathcal{S}.$$

However,

$$\begin{bmatrix} A^T & C^T \end{bmatrix} \left[(0 \times \mathbb{R}^p) \cap \ker \begin{pmatrix} B^T & D^T \end{pmatrix} \right] = C^T \ker D^T.$$

So indeed $C^T \ker D^T \subseteq \mathcal{S}$ for all $\mathcal{S}$ satisfying the weakly unobservable subspace inclusion.

Having proved that $\mathcal{S}^*(A^T, C^T, B^T, D^T) = C^T \ker D^T$, duality 2.20 then implies $\mathcal{V}^* = (C^T \ker D^T)^\perp = C^{-1} \operatorname{im} D$ concluding the proof.

Having now both computed $\mathcal{V}^*$ and $\mathcal{S}^*$ we can derive quite an interesting statement as to their relation with each other.

Lemma 3.3. *$\mathcal{V}^* = (P\mathcal{S}^*)^\perp$ for all $P \in \mathcal{P}$*

Proof. By statement (ii) of Lemma 2.27 and (v) of Lemma 2.26, $P\mathcal{S}^* = P\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + PB \ker D = PB \ker D = C^T \ker D$. Then using statement (iii) of Lemma 2.26 and Lemma 2.7 and 2.4 $(P\mathcal{S}^*)^\perp = (C^T \ker D)^\perp = C^{-1} \operatorname{im} D = \mathcal{V}^*$.

What is more remarkable is that, upon the assumption that P is non-singular, one can actually show that $\mathcal{V}^* = \mathcal{S}^{*\perp}$.

Lemma 3.4. *The controllable weakly unobservable subspace of a passive linear system is given by*

$$\mathcal{R}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle.$$

Proof. By definition

$$\mathcal{R}^* := \mathcal{V}^* \cap \mathcal{S}^*$$

which when plugging in our derived expressions for $\mathcal{V}^*$ and $\mathcal{S}^*$ yields

$$\mathcal{R}^* = (\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D) \cap C^{-1} \operatorname{im} D$$

however, by statement (v) of Lemma 2.27 this implies

$$\mathcal{R}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle$$

concluding the proof.

Lemma 3.5. *The distributionally weakly unobservable subspace of a passive linear system is given by*

$$\mathcal{N}^* = B \ker D + C^{-1} \operatorname{im} D.$$

Proof. Again by definition

$$\mathcal{N}^* := \mathcal{V}^* + \mathcal{S}^*.$$

then substituting the derived expressions for $\mathcal{V}^*$ and $\mathcal{S}^*$ produces

$$\mathcal{N}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle + B \ker D + C^{-1} \operatorname{im} D$$

But note that by statement (iv) of Lemma 2.27 the subspace $\langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle$ is already contained in $C^{-1} \operatorname{im} D$ so we can rewrite the sum more succinctly as

$$\mathcal{N}^* = B \ker D + C^{-1} \operatorname{im} D.$$

Lemma 3.6. *The distributionally weakly unobservable subspace of a passive linear system is equal to the state space.*

$$\mathcal{N}^* = \mathbb{R}^n.$$

Proof. The above lemma is equivalent to the statement that $\mathcal{N}^{*\perp} = 0$, which we will now demonstrate. Using Lemma 2.7 we find the orthogonal complement of our derived expression for $\mathcal{N}^*$ in Lemma 3.5:

$$\begin{aligned}\mathcal{N}^{*\perp} &= (B\ker D + C^{-1}\operatorname{im} D)^{\perp}\\ &= (B\ker D)^{\perp} \cap (C^{-1}\operatorname{im} D)^{\perp}\\ &= (B^T)^{-1}\operatorname{im} D^T \cap C^T \ker D^T.\end{aligned}$$

But by Lemma 2.28 $(B^T)^{-1}\operatorname{im} D^T \cap C^T \ker D^T = 0$ and so $\mathcal{N}^{*\perp} = 0$ and therefore, after taking the orthogonal complement of both sides, we arrive at $\mathcal{N}^* = \mathbb{R}^n$.

It cannot be stated enough how remarkable the results of this section are. Unlike the reachable and controllable subspaces, no closed form representation of $\mathcal{S}^*, \mathcal{V}^*, \mathcal{R}^*, \mathcal{N}^*$ can be achieved by using only the fundamental subspaces of (A, B, C, D) for general linear systems. However, the assumption of passivity not only allows us to achieve such representations, we can also do so by exclusively using $B\ker D$, $C^{-1}\operatorname{im} D$ and A-invariance.

Even more surprisingly, is that $\mathcal{N}^*$ turns out to be the whole state space $\mathbb{R}^n$. One important corollary of this, is that surjectivity of the system when treated as a map from the input space to the output space, turns out to be equivalent to the surjectivity of the matrix map $\begin{pmatrix} C & D \end{pmatrix}$. Additionally, under the assumption that P is nonsingular, injectivity turns out to be equivalent to the injectivity of the matrix map $\begin{pmatrix} B \\ D \end{pmatrix}$.

4 Nine-fold Canonical Decomposition of a Passive Linear System

4.1 Introduction

The following lists a result obtained by H. Aling and J. M. Schumacher (see Aling and Shumacher 1984 p. 792), which will be stated without proof. It is the culmination of the geometric approach to the study of linear systems, and decomposes the system based on its fundamental subspaces.

Theorem 4.1 (Nine-fold Canonical Decomposition)**.** *Given a linear system $\Sigma(A, B, C, D)$ there exists a partition of the state space $\mathbb{R}^n$ into subspaces $\mathcal{X}_i$, the space of inputs $\mathbb{R}^m$ into subspaces $\mathcal{U}_i$, and the space of outputs $\mathbb{R}^p$ into subspaces $\mathcal{Y}_i$ such that the following is true.*

1.

$$\mathbb{R}^n = \bigoplus_{i=1}^{9} \mathcal{X}_i$$
$$\mathbb{R}^m = \mathcal{U}_1 \oplus \mathcal{U}_2$$
$$\mathbb{R}^p = \mathcal{Y}_1 \oplus \mathcal{Y}_2$$

2.

$$\begin{aligned} \langle A \mid \operatorname{im} B\rangle &= \mathcal{X}_3 \oplus \mathcal{X}_4 \oplus \mathcal{X}_6 \oplus \mathcal{X}_7 \oplus \mathcal{X}_8 \oplus \mathcal{X}_9 \\ \langle \ker C \mid A\rangle &= \mathcal{X}_1 \oplus \mathcal{X}_3 \oplus \mathcal{X}_7 \\ \mathcal{S}^* &= \mathcal{X}_7 \oplus \mathcal{X}_8 \oplus \mathcal{X}_9 \\ \mathcal{V}^* &= \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \mathcal{X}_3 \oplus \mathcal{X}_4 \oplus \mathcal{X}_7 \oplus \mathcal{X}_8 \\ B^{-1}\mathcal{V}^* \cap \ker D &= \mathcal{U}_1 \\ C\mathcal{S}^* + \operatorname{im} D &= \mathcal{Y}_2 \end{aligned}$$

3. The equivalent system obtained after an appropriate change of basis is given by

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{ccccccccc:cc} A_{1,1} & A_{1,2} & 0 & 0 & A_{1,5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{2,2} & 0 & 0 & A_{2,5} & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & A_{3,5} & A_{3,6} & 0 & A_{3,8} & A_{3,9} & 0 & B_{3,2} \\ 0 & A_{4,2} & 0 & A_{4,4} & A_{4,5} & A_{4,6} & 0 & A_{4,8} & A_{4,9} & 0 & B_{4,2} \\ 0 & 0 & 0 & 0 & A_{5,5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{6,2} & 0 & A_{6,4} & A_{6,5} & A_{6,6} & 0 & A_{6,8} & A_{6,9} & 0 & B_{6,2} \\ A_{7,1} & A_{7,2} & A_{7,3} & A_{7,4} & A_{7,5} & A_{7,6} & A_{7,7} & A_{7,8} & A_{7,9} & B_{7,1} & B_{7,2} \\ 0 & A_{8,2} & 0 & A_{8,4} & A_{8,5} & A_{8,6} & 0 & A_{8,8} & A_{8,9} & B_{8,1} & B_{8,2} \\ 0 & A_{9,2} & 0 & A_{9,4} & A_{9,5} & A_{9,6} & 0 & A_{9,8} & A_{9,9} & 0 & B_{9,2} \\ \hdashline 0 & 0 & 0 & 0 & C_{1,5} & C_{1,6} & 0 & 0 & 0 & 0 & 0 \\ 0 & C_{2,2} & 0 & C_{2,4} & C_{2,5} & C_{2,6} & 0 & C_{2,8} & C_{2,9} & 0 & D_{2,2} \end{array}\right]$$

where $A_{i,j} : \mathcal{X}_j \to \mathcal{X}_i$, $B_{i,j} : \mathcal{U}_j \to \mathcal{X}_i$, $C_{i,j} : \mathcal{X}_j \to \mathcal{Y}_i$ *and* $D_{i,j} : \mathcal{U}_j \to \mathcal{Y}_i$.

4.2 Characterisation of a passive system

Theorem 4.2. *A passive system admits to a six-fold canonical decomposition as defined above, with* $\mathcal{X}_5, \mathcal{X}_6, \mathcal{X}_8 = 0$ *and* $\mathcal{X}_7 = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D\rangle$. *Additionally,* $\mathcal{U}_1 = \ker C^T \cap \ker D$ *and* $\mathcal{Y}_2 = \operatorname{im} C + \operatorname{im} D$ *and therefore contains the output reachable subspace.*

Proof. We can show $\mathcal{X}_5, \mathcal{X}_6 = 0$ as they are the only two subspaces not present when considering the sum $\mathcal{S}^* + \mathcal{V}^*$. As such $(\mathcal{X}_5 \oplus \mathcal{X}_6) \cap (\mathcal{S}^* + \mathcal{V}^*) = 0$, but by Lemma 3.6 $\mathcal{S}^* + \mathcal{V}^* = \mathbb{R}^n$ and so $(\mathcal{X}_5 \oplus \mathcal{X}_6) \cap \mathbb{R}^n = 0$, implying that $\mathcal{X}_5 \oplus \mathcal{X}_6 = 0$.

To investigate $\mathcal{X}_7$, note that it is the subspace uniquely identified by

$$\begin{aligned} \mathcal{X}_7 &= \mathcal{S}^* \cap \mathcal{V}^* \cap \langle \ker C \mid A \rangle \\ &= \mathcal{R}^* \cap \langle \ker C \mid A \rangle. \end{aligned}$$

However note that $\mathcal{R}^* \subseteq \langle \ker C \mid A \rangle$ by statements (viii) of Lemma 2.26 and (ii) of Lemma 2.27, so

$$\mathcal{X}_7 = \mathcal{R}^* = \langle A \mid B \ker D \cap C^{-1} \operatorname{im} D \rangle.$$

With $\mathcal{X}_7$ defined we can now show $\mathcal{X}_8 = 0$ which follows by first considering that by construction we have $\mathcal{X}_7 \oplus \mathcal{X}_8 = \mathcal{V}^* \cap \mathcal{S}^* = \mathcal{R}^* = \mathcal{X}_7$, as such $\mathcal{X}_8 \subseteq \mathcal{X}_7$, but since by definition $\mathcal{X}_7 \cap \mathcal{X}_8 = 0$, we must necessarily have $\mathcal{X}_8 = 0$.

To show $\mathcal{U}_1 = \ker C^T \cap \ker D$ consider the definition of $\mathcal{U}_1$

$$\mathcal{U}_1 := B^{-1}\mathcal{V}^* \cap \ker D.$$

As such

$$\mu \in \mathcal{U}_1 \Rightarrow CB\mu = Du \text{ and } \mu \subseteq \ker D.$$

Then

$$\begin{aligned} \mu^T CB\mu &= \mu^T Du \\ \mu^T B^T PB\mu &= 0 \\ PB\mu &= 0 \\ C^T\mu &= 0 \\ \mu &\in \ker C^T \end{aligned}$$

and so

$$\mathcal{U}_1 \subseteq \ker C^T \cap \ker D.$$

The reverse inclusion is evident after considering

$$\mu \in \ker C^T \cap \ker D$$

$$
\begin{aligned}
C^T\mu = 0 \Rightarrow PB\mu = 0 \Rightarrow CB\mu = 0 \Rightarrow CB\mu = D(0)\\
B\mu \in C^{-1}\operatorname{im} D\\
B\mu \in \mathcal{V}^*\\
\mu \in B^{-1}\mathcal{V}^*.
\end{aligned}
$$

The proof of $\mathcal{Y}_2 = \operatorname{im} C + \operatorname{im} D$ is a bit more involved and first requires to show that $\mathcal{U}_1 = \mathcal{Y}_2^\perp$.

$$
\begin{aligned}
\mathcal{Y}_2^\perp &= (C\mathcal{S}^* + \operatorname{im} D)^\perp\\
&= (C\mathcal{S})^{*\perp} \cap \operatorname{im} D^\perp\\
&= (C\mathcal{S})^{*\perp} \cap \ker D.
\end{aligned}
$$

Now to investigate $(C\mathcal{S}^*)^\perp$, first note that, since $\langle A \mid B\ker D \cap C^{-1}\operatorname{im} D\rangle \subseteq \ker C$

$$
C\mathcal{S}^* = C(\langle A \mid B\ker D \cap C^{-1}\operatorname{im} D\rangle + B\ker D) = CB\ker D
$$

and therefore

$$
\begin{aligned}
\mathcal{Y}_2^\perp &= (CB\ker D)^\perp \cap \ker D\\
\mathcal{Y}_2^\perp &= (B^TC^T)^{-1}\operatorname{im} D \cap \ker D.
\end{aligned}
$$

Now take an $\mu \in \mathcal{Y}_2^\perp$, by the above $B^TC^T\mu = Du$ and $\mu^T D = 0$, as such

$$
\begin{aligned}
\mu^T B^T C^T \mu = 0\\
\mu^T B^T P B\mu = 0.
\end{aligned}
$$

Using Lemma 2.13 then gives

$$
\begin{aligned}
&PB\mu = 0\\
&B\mu \in \ker P \subseteq \ker C \subseteq \mathcal{V}^*\\
&\mu \in B^{-1}\mathcal{V}^*\\
&\mu \in \mathcal{U}_1.
\end{aligned}
$$

By arbitrary choice, we conclude that $\mathcal{Y}_2^\perp \subseteq \mathcal{U}_1$, to show the reverse inclusion take an arbitrary $\mu \in \mathcal{U}_1$. By properties of $\mathcal{U}_1$, $D\mu = 0$ and $\mu \in B^{-1}(C^{-1}\operatorname{im} D) = (CB)^{-1}\operatorname{im} D$ so we also have that

$$
CB\mu = Du,
$$

additionally for all $y \in \mathcal{Y}_2$, $y = CB\mu_2 + Du_2$, as such

$$\begin{aligned}
\mu^T y &= \mu^T CB\mu_2 + \mu^T Du_2 \\
&= (C^T\mu)^T B\mu_2 \\
&= (PB\mu)^T B\mu_2 \\
&= \mu^T B^T PB\mu_2 \\
&= \mu^T B^T C^T \mu_2 \\
&= (CB\mu)\mu_2 \\
&= u^T D^T \mu \\
&= 0.
\end{aligned}$$

So $\mu \in \mathcal{Y}_2^\perp$, and therefore $\mathcal{U}_1 \subseteq \mathcal{Y}_2^\perp$, which when combined with the previous result demonstrates that indeed $\mathcal{U}_1 = \mathcal{Y}_2^\perp$.

Then the statement we sought out to prove, namely $\mathcal{Y}_2 = \operatorname{im} C + \operatorname{im} D$ is evident, as

$$\begin{aligned}
\mathcal{Y}_2 &= \mathcal{U}_1^\perp \\
\mathcal{Y}_2 &= (\ker C^T \cap \ker D)^\perp \\
\mathcal{Y}_2 &= \ker C^{T\perp} + \ker D^\perp
\end{aligned}$$

which by statement (iii) of Lemma 2.26 and Lemma 2.4 yield

$$\mathcal{Y}_2 = \operatorname{im} C + \operatorname{im} D.$$

4.3 Characterisation of a passive system with positive definite storage function

In the following section we assume along with passivity, the existence of a positive definite P satisfying the passivity inequality. It turns out that such a positive definite matrix existing has multiple consequences.

Firstly, $\mathcal{X}_7 = 0$ and as such we can reduce our previously defined sixfold decomposition to a five-fold decomposition. This follows since, if P is positive definite, $\ker P = 0$ by Lemma 2.13, and $\mathcal{X}_7 = \mathcal{R}^* \subseteq \ker P$ by statement (ii) of Lemma 2.27. Additionally, since both $\mathcal{X}_7$ and $\mathcal{X}_9$ are equal to zero, then $\mathcal{X}_9 = \mathcal{S}^*$ which by statement (ii) of Lemma 2.27 gives $\mathcal{X}_9 = B \ker D$. What is also interesting is that we can derive an even stricter definition for $\mathcal{U}_1$.

Lemma 4.3. $\mathcal{U}_1 = \ker B \cap \ker D$.

Proof. Using statement (v) of Lemma 2.26

$$\begin{aligned} \mu \in \ker C^T \cap \ker D &\iff C^T \mu = 0 \\ &\iff PB\mu = 0 \\ &\iff B\mu = 0 \text{ since } \ker P = 0 \\ &\iff \mu \in \ker B \cap \ker D. \end{aligned}$$

The above effectively means that any $u(t) \in \mathcal{U}_1$ does not affect the behavior of the system, which along with the fact that the output reachable subspace is fully contained in $\mathcal{Y}_2$, allows for the reduction of the input and output space by discarding $\mathcal{U}_1$ and $\mathcal{Y}_1$, without loss of behavior.

Remark 4.4. *A passive system with positive definite storage function admits to a 5 fold canonical decomposition as defined above, with* $\mathcal{X}_5, \mathcal{X}_6, \mathcal{X}_7, \mathcal{X}_8 = 0$ *and* $\mathcal{Y}_1$ *and* $\mathcal{U}_1$ *not playing a role in the dynamics of the system. In appropriate coordinates, we can write the block matrix as:*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{ccccc|c} A_{1,1} & A_{1,2} & 0 & 0 & 0 & 0 \\ 0 & A_{2,2} & 0 & 0 & 0 & 0 \\ A_{3,1} & A_{3,2} & A_{3,3} & A_{3,4} & A_{3,9} & B_{3,2} \\ 0 & A_{4,2} & 0 & A_{4,4} & A_{4,9} & B_{4,2} \\ 0 & A_{9,2} & 0 & A_{9,4} & A_{9,9} & B_{9,2} \\ \hline 0 & C_{2,2} & 0 & C_{2,4} & C_{2,9} & D_{2,2} \end{array}\right]$$

5 Acknowledgements

This thesis could not have been possible without the insight, tutelage, and patience of Professor Kanat Camlibel. I could not express more gratitude towards Professor Camlibel for agreeing to supervise me.

References

Aling, H. and H. J. Shumacher (1984). “A nine-fold canonical decomposition for linear systems”. In: *International Journal of Control* 39.4, pp. 779–805.

Trentelman, Harry L., Anton A. Stoorvogel, and Malo Hautus (2002). *Control theory for linear systems*. University of Groningen.

Willems, Jan C. (1972). “Dissipative dynamical systems part II: linear systems with quadratic supply rates”. In: *Archive for Rational Mechanics and Analysis* 45, pp. 352–393.