On Control Allocation and its Applicability for Dual-Stage Actuator Systems

Andreas Zürcher\textsuperscript{1}, Kristian Dimitrov\textsuperscript{1}, Manuel Schwartz and Sören Hohmann

Abstract—The aim of this paper is to study the applicability of the control allocation (CA) framework for dual-stage actuator (DSA) systems - a subclass of overactuated precise positioning devices, characterised by serial connection of a coarse and fine stage. In contrast to the typical control frameworks developed for DSAs, such as master-slave and decoupled design, the CA framework approaches the control in a modular way, clearly separating the motion control from the allocation of efforts among the actuators. To achieve this, the concept of redundant controllability is introduced with the aim to analyse weakly input redundant LTI systems to determine the virtual effort that enables the separation of the system in terms of the CA approach. As a result, it shows that with serially connected actuators, the redundancy is clearly present in the displacement domain. In addition, the application of the framework is shown and the possible control approaches are discussed.

I. INTRODUCTION

Dual-Stage Actuators (DSA) are mechatronic systems characterized by the serial connection of two complementary actuators, designed for driving precise positioning applications. The first (coarse) stage is designed to cover the total range of motion, and the second (fine) stage provides the required accuracy and bandwidth. Having two independent actuators contributing to the one dimensional output positioning, the DSA falls under the broader category of the overactuated systems.

Overactuation always implies some kind of input redundancy (IR) \[1\]. According to \[2\] this term can be further split into two types, namely strong and weak IR. Depending on where in the system the redundancy occurs, in the first case, the input operator is rank deficient, whereby in the latter case the redundancy occurs somewhere in the system, i.e. the state- or the output space. Upon closer examination of the weak IR, it becomes evident that this characterization encompasses several potential system configurations, as illustrated in Fig. 1. Consequently, the requirements for controlling such systems vary based on the specific internal structure in question, such as the need to compensate for low-level dynamics or implement a decoupling controller.

In all input redundant systems, apart from achieving the standard output control objectives, there arises the need to address an allocation problem. In the context of motion control, this task can be understood as how to distribute the total control effort among the actuators by considering their characteristics and constraints. In order to enable a structured approach for control and allocation design, in the last decades a wide variety of solutions under the term Control Allocation (CA) have been developed. When examining the literature, two primary CA frameworks become evident: Historically the first one to appear is characterized by the extraction of a virtual control effort in the system, whose dimensionality equals the number of controlled system outputs (see \[3\], \[4\], \[5\] and the references therein). This allows a system decomposition into a square subsystem (see Def. 3) and an input redundant part. The advantage of this approach lies in the clear separation of the control and the allocation tasks. All control modules arise as square systems, and only the allocation module has to resolve the overactuation (see Fig. 2). However, its applicability depends on whether such virtual control effort can be found. Introduced by the work of \[2\], another type of control framework under the term control allocation has been established (and well researched, e.g., by \[6\], \[7\], \[8\]). The strategy proposed there solves the allocation problem alongside with the control task inside a single module. Since this is not a strict separation of control and allocation of the control inputs, this approach is not taken into consideration here.

To indicate the problem statement of this paper, a closer examination of the main idea behind DSAs is required. The specific actuator combination of a coarse and fine
stage implies an interconnection of so called complementary\textsuperscript{1} properties (see Table I).

Due to the unique structure of serially connected complementary actuators, a toolbox of DSA-specific control frameworks has arisen including Master-Slave Design, Decoupled Design, and Parallel Design (see [10], [11]). Analysed and developed mainly within the field of Hard Disk Drives (HDD), these frameworks have established themselves as state-of-the-art for DSAs in other applications. Common across all of them is the use of two controllers designed with respect to each stage, and specific feedback connections which lay down the allocation principle between the controlled stages. In a survey paper [10], the advantages and drawbacks of these frameworks are analysed. Irrespective of the framework, both controllers can be designed using a variety of methods, and many recent papers propose new algorithms which provide enhanced performance (e.g., based on sliding mode control [12]).

Nevertheless, a specific connection between the controllers represents a restriction for the possibilities how to distribute the commands among the stages. As [13] argues, specific trajectories are not achievable with these existing schemes. Additionally, the established DSA control strategies approach the typical control objectives (stabilization, tracking, disturbance rejection, etc.) and the effort allocation not in a separate way, but they are a product of the controller tuning and the specific feedback connections. In order to alter the allocation among the controlled actuators, the individual controllers have to be retuned. However, this would in turn affect the control objectives. Furthermore these approaches are not scalable and therefore restricted to the typical DSA setting.

In the paper at hand, the assumption is made that the targeted exploitation of the complementary properties is closely associated to the distribution of redundant inputs and therefore should, as far as possible, be treated in a separate module, independent of the controller design. Having the advantages of modular design and a variety of developed allocation schemes, the CA framework represents a promising approach to address the mentioned issues. In the last decade the CA framework has found increasing application in many engineering fields (power electronics [14], power systems [15], precise positioning devices [16] and nuclear fusion technologies [17]), but still its utilisation for DSA systems has been insufficiently researched.

Hence the novelty of this paper is to analyse the capability of the CA framework to DSA. To do this, a general analysis of the redundancy within weakly IR systems is given. The paper is organised in the following way: Section II provides formal definitions of overactuation, input redundancy, as well as a brief introduction to the CA framework. In section III the concept of redundant controllability is given. Furthermore the existence of the virtual effort in weakly IR systems is investigated. In Section IV the DSA system is analysed with respect to the CA framework requirements. In Section V, a CA framework is formulated for DSAs.

### II. Background

#### Nomenclature

Within this paper calligraphic symbols $\mathcal{C}$ are used to represent operators which can be understood as a generic control law. $\text{im}(\cdot)$ denotes the image of a matrix. $\ker(\cdot)$ denote the right nullspace of a matrix. $\Sigma_{\cdot,\cdot}$ denote a LTI system according to Eq. 1 and $\mathcal{X}_{(\cdot)}$ is the corresponding state space.

#### A. Overactuation vs. Input Redundancy

In order to give a delineation of the terms according to the current state-of-the-art, we consider the LTI System $\Sigma$:

$$
\begin{align*}
\dot{x}(t) &= A x(t) + B u(t) \\
y(t) &= C x(t),
\end{align*}
$$

with the states $x(t) \in \mathbb{R}^n$, the control inputs $u(t) \in \mathbb{R}^p$ and the systems outputs $y(t) \in \mathbb{R}^q$ at time $t$.

**Definition 1.** (Overactuation) According to [8], $\Sigma$ is said to be overactuated if $p > q$ and $\text{rank}(B) \geq q$ holds. In Addition it is assumed that the system is minimal and $\text{rank}(C) = q$.

A closer look at Def. 1 shows that a case distinction regarding the rank of linear map $B$ can be made, which leads to the definitions of strong and weak IR:

**Definition 2.** (Input Redundancy) An overactuated system is called strongly IR, if $p > \text{rank}(B) = q$ holds, and weakly IR, if $p = \text{rank}(B) > q$ is given.

According to our distinction the CA framework is characterised by the fact, that the control law - irrespective if it is open or closed-loop - has to face additionally the actuator coordination. This leads to the fact, that it has to handle the computation of a higher dimensional signal out of a lower dimensionality. In order to ensure generality, we will indicate this control law as a non-square\textsuperscript{2} operator $\mathcal{C}$:

**Definition 3.** (Square Systems) A system is called square if the number of inputs $p$ is equal to the number of outputs $q$. In an analogous way we define an operator $\mathcal{O} : X \rightarrow Y$ as a square operator, if it is a bijection.

\textsuperscript{1}These properties are consistent with the antagonisms, defined in [9].

\textsuperscript{2}As an example for a square control law, one can think of a classical PID controller. In contrast, Linear Quadratic Regulators (LQR) or Model Predictive Controllers (MPC) are able to handle overactuation, and therefore can be implemented as non-square operators.
B. Control Allocation (CA)

The CA framework is usually divided into three levels as a hierarchical structure. The high-level controller $C_H$, an allocator $C_A$ and possible low-level controllers $C_i$ (see Fig. 2). The proposed division of the control requires separation of the system, i.e., it must be possible to separate the system dynamics, which allows to extract the redundant controllable states. Provided that these states are unique, the “location” where the overactuation starts in the system can be found and specifically addressed by the allocator. This new system input is referred to as a virtual effort (or task) $\tau$ and can usually be understood as the sum of all actuator forces, although an alternative is shown in this paper (see Sec. IV). Remaining dynamics are assigned to the individual input paths and taken into account by the respective low-level control. Thus, all controlled sub-systems appear to be square.

Obviously, this definition cannot be applied to weakly IR without further elaboration. Consider the different system configurations shown in Figure 1, which are united under the term weakly IR. For system II.) and III.), for example, it is possible to find several matrices that satisfies definition 4. In other words there are various subsystems, which have the property of being strongly IR. In order to analyse the separability of overactuated systems in the general case, in the following the concept of selective controllability by [20] is revisited and extended to the concept redundant controllability.

III. Redundant Controllability

**Proposition 1.** Consider the Hermite controllability matrix as introduced in [21] by the matrix pair $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times p}$:

$$H(A, B) := [b_1 \ A_{n-1} b_1 \ b_p \ A_{n-1} b_p] \quad (2)$$

where $b_{(-)}$ is the respective column vector of $B$. $H(A, B)$ can be rewritten as a block matrix in the form of

$$H(A, B) =: \hat{H}(Q_{c,j})_{j \in \{1,...,p\}} = [Q_{c,1} \ ... \ Q_{c,p}], \quad (3)$$

whereby the respective elements of $\hat{H}$ are understood as the selective controllability matrices:

$$Q_{c,j} := [b_j \ A_{n-1} b_j] \in \mathbb{R}^{n \times n}. \quad (4)$$

Based on these matrices, the Kalman criterion can be used to prove complete controllability determined by the j-th input. In the case of full controllability, the question of selective controllability is trivial, which is why it is assumed that this is not the case, i.e. rank($Q_{c,j}$) $< n$. Furthermore, there is no restriction on the algebraic multiplicity of the eigenvalues of $A$. Also, if only the i-th row vector of $Q_{c,j}$ is taken into account, the i-th component of $x(t)$ is selectively controllable by the j-th component of $u(t)$ if, and only if, it contains at least one non-zero element. (Proof see [20, p. 451].)

**Definition 5.** (Redundant Controllability) A component $x_i(t) \in \mathbb{R}$ of the state vector $x(t)$ from the system $\Sigma$ is called redundant controllable if there exists at least two components $(u_j(t)$ and $u_k(t)) \in \mathbb{R}$ with $j \neq k$, of the control vector $u(t)$, for which $x_i(t)$ is controllable, in the sense that there exists a finite time $t_E$ with the corresponding state $x_i(t_E) = x_{i,E}$ so that the initial state $x_i(t_0) = x_{i,0}$ can be transferred into $x_{i,E}$ by both $u_j$ and $u_k$. In other words, $x_i(t)$ must be selective controllable by at least two distinct components from $u(t)$.

**A. Redundant Controllable Subspace**

**Definition 6.** (Redundant Controllable Subspace) The Subspace of $\mathbb{X}$ which is spanned by the redundant controllable components of $x(t)$ is called the redundant controllable subspace, denoted by $\mathbb{X}_R$. The orthogonal complement of $\mathbb{X}_R$ is called the non-redundant controllable subspace $\mathbb{X}_{NR} := \mathbb{X} \setminus \mathbb{X}_R$. Under the assumption that $\Sigma$ is minimal, it holds that $\mathbb{X}_R \oplus \mathbb{X}_{NR} = \mathbb{X}$.
Proposition 2. Consider $X$ as the corresponding state space to $\Sigma$ and $Q_{c,j}$ as the selective controllability matrix as introduced in Prop. 1. By

$$X_{c,j} := \text{im}(Q_{c,j}) \subseteq X$$

the selective controllable subspace $X_{c,j}$ can be defined. From this, the redundant controllable subspace is described through the intersection of the selective controllable subspaces:

$$X_R := \bigcap_{i=1}^{p} X_{c,j}.$$  

∀ $X_R \neq \{0\}$ one can define a according subsystem $\Sigma_R$ by the tuple $(A_R, B_R)$ with the corresponding state vector $x_R(t) \in X_R$. This results in the overall system description by

$$\dot{x}(t) = \begin{bmatrix} A_{NR} & 0 \\ A_t & A_R \end{bmatrix} \dot{x}(t) + \begin{bmatrix} B_{NR} \\ B_R \end{bmatrix} u(t).$$

with the transformed state vector $\dot{x} = \begin{bmatrix} x_{NR} \\ x_t \end{bmatrix}^T = T_{Rx}x$. The matrices $A_{NR} \in \mathbb{R}^{\dim(X_{NR}) \times \dim(X_{NR})}$ and $A_t \in \mathbb{R}^{\dim(X_t) \times \dim(X_t)}$ represents the eigen dynamics of the respective states. Whereas $A_t \in \mathbb{R}^{\dim(X_t) \times \dim(X_t)}$ describes the coupling between these two subsystems. The transformation matrix $T_R$ is given by a basis $(b_i \in \mathbb{R}^{\dim(X_t)})$, $i \in \{1,...,\dim(X_t)\}$ of $X_R$ an orthogonal complements of $b_i$.

Based on the characteristics of $X_R$, respective $A_R$, it is now possible to analyse the separability of the system in terms of the applicability of Control Allocation. More precisely, to investigate the existence of an virtual effort in weakly IR systems, which is done in the following theorem:

Theorem 1. Consider the redundant controllable subspace $X_R$ and the transformed system $\Sigma_R$ as introduced in Prop. 2. Furthermore $\Sigma$ is assumed to be weakly IR. Depending on the size of $X_R$ one can carry out a case distinction:

I. $X_R = \{0\} \Rightarrow \ker(C) \neq \{0\}$: This case implies that the virtual effort can be factorized out of the output matrix $C$. Namely, the redundancy of the system starts in the output.

Proof: Since $X_R = \{0\}$, all $\text{im}(Q_{c,j})$ are linearly independent, thus every input $j$ steers an independent subsystem, which is fully observable. Therefore $C$ contains at least $p$-elements in the form of $C = \{c_j\}_{j \in \{1,...,p\}} \in \mathbb{R}^{\dim(X_t)} \times \mathbb{R}^{\dim(Q_{c,j})}$. With $p > q$ it follows that $\ker(C) \neq \{0\}$ holds.

II. $X_R \subset X \Rightarrow \ker([a_{1,1}^T \ b_{R,1}^T]) \neq \{0\}$: Consider $a_{1,1}^T$ and $b_{R,1}^T$ the first row vector of $A_1$ and $B_R$, respectively. The redundancy starts at the first redundant state of $X_R$, i.e. in the input space of $\Sigma_R$. Depending on whether $\Sigma_R$ is weakly- or strongly IR, there are more than one possible virtual efforts in the system. In the latter case it is unique, whereby in each case a factorization out of $[a_{1,1}^T \ b_{R,1}^T]$ is possible.

Proof: Irrepective of $\Sigma_R$ is weakly- or strongly IR, the number of independent inputs in $\Sigma_R$ is larger than $q$. Thus the number of non-zero elements $p^*$ in $[a_{1,1}^T \ b_{R,1}^T] \in \mathbb{R}^{\dim(X_t)}$ is also larger than $q$. Since $\Sigma$ is observable by definition, $q$ hat to be larger than 1 and therefore $\ker([a_{1,1}^T \ b_{R,1}^T])$ always holds.

III. $X_R = X$: The whole state space is redundant controllable. Depending on the nullspace of $C$ a further case distinction is needed:

III.a $\ker(C) \neq \{0\}$: Analogue to case I., $\tau$ can be factorised out of $C$. However, similarly to case II. (and $\Sigma_R$ is weakly IR) this might not be a unique solution.

III.b $\ker(C) = \{0\}$: Since the nullspace of $C$ is empty, not even a static subsystem can be separated. Thus, without further adjustments (e.g. through the implementation of a decoupling controller), Control Allocation cannot be applied.

Remark 1. The analysis presented focusses only on a single separation of the system. Especially in case II., if $X_R$ is weakly IR and case III. of Th. 1. Deeper considerations of the redundant controllable subspace, as well as the proof of Prop. 2 and Th. 1, case III. are left out for further research.

IV. Dual-Stage Actuator Analysis

For the following analysis a generic DSA configuration is considered, consisting of a linear drive (e.g., spindle motor) as the coarse stage and a piezoelectric actuator as the fine stage. Developing the dynamic physical model of mechatronic actuators is usually approached with lumped parameter models, which include both the electrical and mechanical components that constitute the system. Fig. 3 shows the lumped parameter model of the serially connected actuators (i.e., the DSA). Apart from the represented linear elements, these systems exhibit pronounced non-linear behaviour too, such as hysteresis and stick-slip effect.

Fig. 3. Lumped-parameter schematic of a DSA, separated into fine and coarse stage. I.) Electrical circuits. II.) Mechanical circuit of the whole serial coupled DSA.
Ubiquitous in the DSA literature are certain simplifications of this physical model. First, as stated for example in [22] and [16], current (or voltage) control of the actuators is usually assumed which compensates the back-emf forces. Consequently, the actuators are viewed as pure torque (or force) generators, and the system is modelled only with its mechanical subsystem, with a proportional relationship between the input voltage and the electrical torque assumed. The system is described by the equations of motion, which leads to the following state space representation $\Sigma_D$:

$$\dot{x}_D = \begin{bmatrix} a_{11} & a_{13} & a_{14} \\ b_{11} & b_{12} \\ 0 & 0 & 1 \\ 0 & 0 \end{bmatrix} x_D + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_D \quad (8)$$

and

$$y_D = \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} x_D, \quad (9)$$

where $x_D = [y_m \ y_m \ (y_{pe} - y_m) \ (y_{pe} - y_m)]^T$ are the respective states and $u_D = [F_m \ F_{pe}]^T$ represents the input forces generated by the actuators. The subscript $(\cdot)_D$ stands for DSA. Since only the structure of the matrices are relevant, a detailed description of the individual parameters is omitted. Due to the coupling of the actuators, when a driving force is applied on one stage, its displacement induces a force on the other stage, which alters its state as well. A second very common simplification is to represent the system as a SISO structure, which enables a wide range of choices for the specific control algorithms and the tuning of the parameters to zero, the decoupled system $\Sigma_D^*$ results with the triplet $(A_D^*, B_D^*, C_D^*)$. Irrespective of the simplifications assumed, the DSA systems are overactuated in all cases, as they have two inputs and one output ($p > q \land \text{rank}(C_D) = q$). Furthermore, they are weakly input redundant, as the input matrix is at full rank ($p = \text{rank}(B_D) = \text{rank}(B_D^*)$), which is consistent with Def. 2. As introduced in Sec. II-B, for the CA framework it is required to define a virtual effort $\tau$. In the domain of mechatronic positioning systems, the virtual effort is exclusively interpreted as forces or moments. By following this interpretation, one can distinguish between several forces occurring in the system. Intuitively, it seems first to choose the electrical forces that are generated by the actuators and act as input to the system. However, this approach fails for both the coupled and the decoupled system ($\Sigma_D$ and $\Sigma_D^*$). This is because both forces act on different masses in the system. In fact, neither the electrically generated forces nor any other force occurring in the system can be understood as the beginning of redundancy. This is only the case with actuators that are connected in parallel.

Considering $\Sigma_D^*$ as the actuator model, one can apply case I. of Th. 1, which means that $\mathcal{X}_{RD}^* = \{0\}$ holds. This allows to factorise $\tau$ out of the output matrix $C$. If, on the other hand, we look at the fully coupled system $\Sigma_D$ the complete state space became redundant, which is consistent with Th. 1, case III.a. Therefore, defining $\tau$ as the output of the system is also a valid solution. Nevertheless, due to the coupling of the actuators, this may not be a unique solution. In summary, the preceding analysis shows that with a serial connection of actuators, the redundancy certainly occurs at the output of the system, i.e. the total displacement:

$$\tau = \sum_i y_i = y_m + y_{pe} = M^T Y, \quad (10)$$

where $M$ becomes the identity $I \in \mathbb{R}^{1 \times 2}$ and $Y = [y_m \ y_{pe}]^T$ represents the vector of actuator displacements. The result is also consistent with a physical interpretation of the serial connection. In this, the actuators are interpreted as springs connected in series, whose spring force is dependent on the electrical actuation.

V. CONTROL ALLOCATION FORMULATION FOR DUAL-STAGE ACTUATORS

Fig. 4 illustrates the resulting control allocation structure which follows from the proposed interpretation.

![Diagram](image-url)

Fig. 4. Suggested CA framework for a DSA system. $\Sigma_{(\cdot)}$ represents the whole decoupled dynamics of the coarse- and fine stage. $C_{L(\cdot)}$ are the respective low-level controllers.

Representing the virtual effort in the position domain, and more specifically as the output position of the DSA, leaves the high-level system to have the trivial description $\tau = y$. Thus, a high-level system essentially drops, and the low-level system is the whole Dual-Stage Actuator. In this sense, the resulting low-level and high-level systems are decoupled. The mapping $M$ constitutes a simple static addition of the displacement of each stage. The dynamic behavior of the DSA is modelled in the subsystem between the system inputs $\bar{u}_{cs}$ and $\bar{u}_{fs}$, and the displacements $y_{cs}$ and $y_{fs}$. Here a model with actuator interaction can be developed (e.g. $\Sigma_D$), or two independent ones (e.g. $\Sigma_D^*$), as described in Sec. IV. Essential for the control framework is the square structure of this low-level system, having two inputs and two outputs.

Individual motion controllers $C_{Lcs}$ and $C_{Lfs}$ can be developed independently for each actuator, without considerations of the over-actuation present in the system. Both controllers have a SISO structure, which enables a wide range of choices for the specific control algorithms and the tuning.
approach can be based on typical control systems theory. For models with actuator interaction, the low-level control shown in Fig. 4, can be extended by a decoupling network as realised in [22]. The essential task of the low-level control is the stabilisation of the actuator dynamics, and the targeted consideration of robustness and response time criteria.

The allocation algorithm is located in a separate controller unit, responsible for distributing the reference trajectory among the actuators. This allows the realisation of wide range of allocation schemes and the targeted reconfigurability without retuning the separately developed motion controllers. As in typical control allocation applications, the distribution is based on the constraints (rate and magnitude) of the actuators and their controlled transient response. Characteristic for the dual-stage actuators is that the coarse stage exhibits a considerably lower mechanical bandwidth in comparison to the fine stage. As [23] points out, neglecting actuator dynamics, especially when one or more actuators have low bandwidth, can negatively impact overall system behaviour and potentially lead to instability. The significant difference of the controlled actuator transient responses necessitates an allocation algorithm which considers them. Such algorithms are commonly developed by introducing dynamic compensations in the allocation unit, and thus usually denoted as dynamic control allocators.

Further research of the modern methods for DSA control, a developed framework was found which is consistent with the proposed framework in this section. The work in [13] proposes a control structure, in which the reference trajectory is distributed using signal processing algorithm (spatial-temporal filters). Then the distributed trajectories are applied to inversion-based feedforward filters, compensating the following controlled actuators. The compensated signals are finally fed in individual motion controllers, developed individually for each stage. The assumptions for decoupled actuator dynamics and linear behaviour are also met. Without approaching the DSA control problem from the perspective of the control allocation framework, the paper has presented a concept congruent with it.

VI. CONCLUSIONS

This paper uses the introduced concept of the redundant controllable subspace (stated in Th. 1) to study the separability of weakly IR systems, into to fully redundant controllable- and non-redundant controllable states. At this basis, the applicability of the CA for DSA is analysed and as a result, the need for a high-level controller no longer exists. The aim of the outlook is to examine the suitability of the CA setting specifically in relation to the complementary properties of the DSA (see Table I). Furthermore, improvements of Th. 1 are sought, especially in the direction of fully coupled- or nonlinear systems.

REFERENCES