PhD program Technology, Innovation and Management (TIM)
XXXIII Cycle
University of Bergamo and University of Naples Federico II

Ph.D. Dissertation

Implementation of biomimetic principles in optimization tools for design

Author:
Antonio Caputi
Mat. 1055090

Supervisor:
Prof. Davide Russo

Abstract

In the past years, many engineers and researches have focused the attention on the biological systems: Biomimetics is the name of the discipline which studies the formulation of technical solutions inspired by the observation of living beings. This particular research field has been widely developed, and the reason for such interest is easily explained by the extraordinary performance of the structures and systems which may be found in nature.

This dissertation focuses on the application of Biomimetic principles in the creation of tools for the design. These bio-inspired paradigms are ideas expressed at a high conceptual level, and their implementation for practical applications requires an appropriate methodological framework. In this work, the attention has been focused on understanding the mechanisms that allow the living being to have high performances, and implementing this knowledge for technical purposes.

As a unifying procedure, this research proposes the formulation of optimization problems that are able to translate the strategies used by the living beings in order to achieve different tasks. More in detail, constrained optimization procedures are particularly suitable for the creation of a general schema since the objective functions are able to model the tasks that must be fulfilled by the system, and the constraints are the mathematical expression of the particular context, or rather the physical laws governing the system itself.

Four biomimetic paradigms have been taken in account: the hierarchical multilevel organization of the biological tissues, the directional anisotropy of the living matter, the redundancy of the limbs, the mechanical flexibility of the organic structures. Starting from this conceptual ideas, different applications have been developed in order to test the effectiveness of the methodology.

The first example is the proposal of a modified BESO topology optimization approach, which takes in account the configuration of the stress field, and in particular the slope of the principal directions of the tensor field. Based on this analysis, a procedure is described which includes the definition of a new mesh able to better represent the sub-structures that naturally rise in the optimization process.

As a second case study, it had been described the implementation of an optimization procedure applied to a redundant kinematic structure. The main idea consists in introducing a redundant kinematic couple in the architecture of an existing machine. This modify causes the mathematical indeterminacy of the solution of the inverse kinematic problem, and the consequent possibility to fulfil a second task. The secondary task can be expressed by a wide variety of equations, which may be adopted as objective functions of a constrained optimization problem. The solution of the optimization problem leads to an improvement of specific indexes which characterize the performance of the mechanism.

A further example is the proposal of an ontological framework for the analysis and synthesis of compliant mechanisms. The motivation of the proposed methodology stands in the effort to get the best from two points of view: the design of compliant mechanisms by the mean of continuum topology optimization, and the use of a discrete approach. Starting from a functional requirement, expressed in terms of relative values of some elements of the stiffness matrix of the structure, it is possible to obtain the topologies fulfilling such specifications. the presented results are mostly conceptual: actually, the schema is a starting point for the design of a synthesis tool, and the idea of hierarchical organization of the structures have to be formalized and integrated in actual algorithms, and, finally, in numerical procedures.

Sommario

Negli ultimi anni, molti ingegneri e ricercatori hanno focalizzato la loro attenzione sui sistemi biologici: biomimetica è il nome della disciplina che si occupa della formulazione di soluzioni tecniche ispirata dall'osservazione degli esseri viventi. Questo particolare campo di ricerca è stato ampiamente sviluppato, e tale interesse può essere facilmente spiegato dalle straordinarie performance dei sistemi e delle strutture che si possono trovare in natura.

Questa tesi si focalizza sull'applicazione dei principi biomimetici nella creazione di strumenti per la progettazione. Questi paradigmi bio-ispirati sono idee espresse in forma concettuale, e la loro implementazione in applicazioni pratiche richiede un adeguato schema metodologico. In questo lavoro, l'attenzione è stata posta nella comprensione dei

meccanismi che permettono agli esseri viventi di avere straordinarie capacità, e applicare questa conoscenza nella tecnica.

Come procedimento generale, questa ricerca propone la formulazione di specifici problemi di ottimizzazione atti a tradurre le strategie usate dagli esseri viventi al fine di raggiungere i loro obiettivi. Più precisamente, le procedure di ottimizzazione vincolata sono particolarmente adatte alla creazione di uno schema di base, in quanto, nella loro forma generale, le funzioni obiettivo possono modellare il compito che deve essere svolto dal sistema, e le equazioni di vincolo possono essere considerate l'espressione matematica del particolare contesto, ovvero le leggi che governano il sistema stesso.

Quattro paradigmi biomimetici sono stati considerati: l'organizzazione gerarchica multilivello dei tessuti biologici, l'anisotropia direzionale della materia vivente, la ridondanza degli organi, la flessibilità meccanica delle strutture organiche. Partendo da queste idee concettuali, sono state create diverse applicazioni al fine di verificare la validità della metodologia.

Un primo esempio è la proposta di un approccio all'ottimizzazione topologica basata su una versione modificata del BESO (Bidirectional Evolutionary Structural Optimization), il quale tiene conto del campo tensoriale che rappresenta lo stato di tensione all'interno della struttura, e, in particolare, delle direzioni principali. Viene quindi descritta una procedura basata su tale analisi, che include la definizione di una mesh di elementi alternativa per la modellazione della struttura, e che meglio è capace di rappresentare le sottostrutture che vengono generate dal processo di ottimizzazione.

Come secondo caso studio, viene descritta l'applicazione di una procedura di ottimizzazione a un meccanismo dotato di ridondanza cinematica. L'idea di fondo è quella di introdurre un grado di libertà ridondante nell'architettura di un meccanismo esistente, come, ad esempio, una macchina utensile. Tale modifica causa la nascita di una indeterminazione matematica della soluzione del problema di cinematica inversa, e la conseguente possibilità di far svolgere al meccanismo un compito secondario. Questo compito secondario può essere definito da una vasta gamma di equazioni, che possono assegnate come funzioni obiettivo per il problema di ottimizzazione. La soluzione di quest'ultimo comporta il miglioramento delle prestazioni del meccanismo.

Un ulteriore esempio è rappresentato dalla proposta di uno schema ontologico per l'analisi e la sintesi di meccanismi cosiddetti "compliant" (flessibili). L'origine della

proposta sta nel tentativo di prendere il meglio da due diversi approcci al problema: il design dei meccanismi “compliant” tramite l’utilizzo di tecniche di ottimizzazione topologica, o, come alternativa, l’utilizzo di un approccio che adotta elementi discreti. Partendo dalle richieste funzionali, espresse in termini di valori relativi degli elementi della matrice di rigidezza, è possibile ottenere strutture che effettivamente soddisfano tali requisiti. I risultati presentati sono principalmente concettuali: lo schema proposto è un punto di partenza per l’analisi e la progettazione di meccanismi “compliant”, e l’idea dell’organizzazione gerarchica delle strutture deve essere formalizzata e integrata in un algoritmo, e in una effettiva procedura numerica.

List of Acronyms

BESO	Bidirectional Evolutionary Structural Optimization
CM	Compliant Mechanism
CNC	Computer Numeric Control
CONLIN	Convex Linearization
ESO	Evolutionary Structural Optimization
FEM	Finite Element Method
GCMMA	Globally Convergent Method of Moving Asymptotes
KS	Kinetic-Static
KKT	Karush Kuhn Tucker
LEM	Lamina emergent mechanism
MMA	Method of Moving Asymptotes
OC	Optimality Criteria
PDE	Partial Differential Equation
PRBM	Pseudo Rigid Body Model
QP	Quadratic Programming
RC	Rejection Criterion
RR	Rejection Ratio
SIMP	Solid Isotropic Material with Penalization
SQP	Sequential Quadratic Programming

Summary

1 Biomimetic approach to design and production

1.1 Scope of the research

The technologies involved in the implementation of the digital factory and the new productive processes, such as additive manufacturing, are changing our idea of information exchange, resources management, and goods production. The main trend in industry 4.0 revolution is the creation of systems and processes that are integrated and interconnected, in which every single subject, part, and component is aware of the state of all the other elements, in order to better accomplish its tasks. In this scenario, the paradigm “keep it simple to make it robust” will no longer work, and this trend can be resumed in two words: increase complexity. Therefore, scientists, researchers and engineers have been called upon to find new approaches to generate better and better technical solutions, in order to deal with complex systems.

In human history, one of the main inspirations for the solution of complex technical problems has been the nature. As it will be highlighted in chapter 2 of this work, in the past years, many have focused the attention on the biological systems: the reason of such interest is easily explained by the extraordinary performances of structures which may be observed studying animals and plants. A vast literature has been produced about this topic, even if spread under different labels (biomimetic, bionics, biomimesis, biomimicry, bionics, biognosis, biologically inspired design).

The main goal of this research is the development of innovative tools for the design, inspired by nature indeed. Software applications (such as CAD applications, FEM software, simulation software, programs for topology optimization), strategies for the integrated design (structure and control developed at the same time), design principles, guidelines and best practices, are all considered design tools. This topic is relevant because design is a fundamental part of the productive process, and this is true for the design of production means, and for the product design as well. The original contribution of this research consists in identifying working principles which can be integrated in design tools: such principles are alternative to the ones generally adopted in commercial software and in the common engineering practices.

1.2 Outlook of the research

The proposed approach for the innovation of the design procedures can be resumed, as a general outline, in three main stages. As a first step, some features characterizing living beings will be identified, and this will be done starting from the analysis of scientific literature caried out in chapter 2. The four main principles identified are reported in the upper part of figure 1.1: the multi-level hierarchical organization of structures, the directional anisotropy of the living matter, the redundancy of biological systems, the compliant behaviour of the biological structures.

As a second stage, high level conceptual ideas have been implemented in different computational means, in order to allow quantitative evaluations, and measure the effectiveness of the method. In practice, this is done by modifying the theoretical models used in some common design tools, according to the biomimetic principles. For instance, in order to use the directionality concept, the topology optimization of a structures should be carried out on the base of finite element analysis adopting a mesh of elements which are not arranged in a regular grid of identical squares, as it is done in most standard synthesis procedures: on the contrary, the elements are modified according to the slope of the main directions of the stress tensor field instead. Another example is the design and the control of the kinematic chain of a machine tool, where the architecture is modified in order to have kinematic redundancy, allowing the optimization of the machine performances. The lower part of figure 1.1 provides a graphical resume of some outcomes in this applicative phase.

Finally, a third stage of the research is the application of the revised design tools. Actually, as a last resort, the adoption of models and methodologies to solve actual technical issues is the final scope of engineering activity, and, for this reason, it is necessary providing some applicative case studies, when possible. Furthermore, the numerical outcomes obtained by the application of the new processes may be confronted with the standard technical solutions, providing an objective validation of the methodology.

1.3 Pillars of the research compared to the existing state of the art, and original contribution

Biomimetic is not a new concept, and a great number of researches have dealt with it in the past. Despite all the efforts, successful biomimetic solutions are not so common: this is mainly due by the low efficiency in operating an effective transfer of the working principles of biological structures in the design of human artefacts. In general, with remarkable exceptions, the approach in realizing a biomimetic application consists in mimicking a particular aspect of an organism, such as the shape, or a particular movement, in order to replicate its performances or behaviour. Many researches showed that this top-down strategy doesn't work in most of the cases. This is mainly due by the fact that living beings have generative principles that are very different with respect to the manufacturing processes. For this reason, simply replying, for instance, the shape of a certain biological structure, does not ensure to obtain the same functionalities: this is mainly caused by the fact that this approach does not consider all the different levels in which all these systems are organized.

On the contrary, the basic idea adopted here is the development of a bottom-up strategy: this means that the goal of the proposed biomimetic framework, is not reproducing a certain shape, structure or layout, but recognizing the principles that generated a particular solution, and apply them in the design process. The mathematical tool used in this work, for bridging the conceptual ideas extracted by the observation of nature, is the optimization. In particular, constrained optimization will be used in the next chapters (chapter 3 and chapter 4): also called mathematical programming, it is a suitable procedure for the purpose, because it involves objective function, which depends by the functional requirement fulfilled by the observed organism, and constraint functions, which depend by the particular investigation field. This should allow to carry out a sort of technology transfer, from the field of biology to the field of the technique.

Reassuming, the main pillars of the research are basically two:

- find technical solutions inspired by some principles derived by the observation of nature;

- create a design tool or procedure by the application of an optimization process to a bio inspired principle.

When possible, such as for the redundancy in chapter 4, the optimization problem will be defined and solved, and the solution of the actual design problem is presented. In other cases, such as the hierarchical multilevel organization of the structures in chapter 2, a more general framework will be presented. This ontological base will be used in order to formulate an optimization problem, such as in chapter 5, or for modifying an already existing optimization procedure, such as in chapter 3. Further details about the said biomimetic concepts, the related state of the art, and the proposed implementation in actual design tools are resumed below.

1.3.1 Hierarchical multilevel structures

Multilevel organization of the living matter is the first principle taken in account. Chapter 2 will provide an overview of biomimetic researches inspired by hierarchical bio-structures. Many works are applied researches, mostly related to the study of new materials. Other authors dedicated to early stages of the design activity, trying to use ideas inspired by organism at a higher conceptual level.

The causality model SAPPhYRE, and the knowledge-based design environment DANE, are two examples of conceptual tools: their purpose is the transfer of the knowledge from the biology domain to the technical domain. In this kind of applications, the domain passage is provided using the functional analysis, which provides an ontological framework necessary to manage the information extracted by the observation of the living beings.

Even if the functional schemes have been used to support the comprehension of the engineers regarding the biological mechanisms, still the synthesis of the solution is due to the creativity of the designers. On the contrary, an ontological framework is introduced in the first stage of the present research, which has the purpose of providing a support to the creation of automatic tools for the synthesis of technical solutions.

1.3.2 Directional anisotropy

The optimization of structures design is complex task, and represent an important topic for engineering. As it will be explained in chapter 3, one of the main difficulties is that Topology Optimization methodologies are the synthesis of results coming from different knowledge areas, such as engineering, mathematics and informatics.

Many researchers have contributed providing a number of outstanding works, describing different approaches: the one based on SIMP material interpolation scheme, and the evolutive approaches as ESO and BESO, are some of the main paradigms applied in actual, and even commercial, applications. Anyway, as it will be shown later on, these kinds of implementations are affected by some limitations.

A proposal of modification of the well-established topology optimization methodologies will be presented in chapter 3. More precisely, it will be proposed the adoption of a different strategy for the definition of the finite element discretization mesh. The idea is that the discretization of the continuum could be done according to the stress state of the optimized structure, in order to improve the performance in terms of volume fraction and total strain energy limitation or, equivalently, compliance limitation. The implementation of a directional anisotropy to topology optimization is an application of multilevel approach to the solution of actual technical problems.

1.3.3 Redundancy

Redundancy is a typical biological feature: organisms are characterized by redundancy in terms of number of organs for instance, or are provided by limbs able to fulfil different tasks at the same time. In this way, animals and plants are able to adapt to the surrounding environment in a more efficient way. Such adaptability is an appealing feature for technical systems too, and the study of redundant systems is a common topic in many fields.

One of the most representative examples of this trend is the study of robot manipulators: since robotics has been one of the first research fields where bio-inspired application appeared, redundancy of kinematic chain of the manipulators is a well-established topic,

in which elegant and effective theoretical solutions have been formulated regarding the control theory.

In chapter 4, the idea of redundancy will be revised, and in particular a novel approach to concurrent structure and control design will be proposed. The outcome of the proposed design approach is a potential improvement of the performance of production means such as milling machines and 3D printers for rapid prototyping.

1.3.4 Compliant structures and mechanisms

Design of compliant mechanism is another interesting topic, which have some common aspects to the optimization of structures, and design of mechanism indeed. It is a frontier research field, and, in the last years, many works have been proposed: despite the fact that many of them are interesting in terms of methodologies and results, the main drawback in approaching this particular argument is the lack of standardization. In fact, it is possible to read outstanding researches about origami inspired robots, manipulators based on notch hinges, monolithic parallel mechanisms: unfortunately, every single approach is presented referring to a particular model framework. In other words, confronting different results from different researches is really difficult because the hypothesis and assumptions are different.

Chapter 5 is dedicated to the description of a first draft of a possible base for the construction of an ontology for the description of compliant mechanisms. The main idea is providing a classification for the elements belonging the compliant structures, for the models describing the behaviour of such elements, and for the functional requirements. According to the hierarchical organization of the systems described in chapter 2, a key aspect of the ontological framework is the possibility to model an element as a substructure, which allows to provide a multilevel description of the system. Furthermore, it will be provided an example of application of the ontological framework to the actual synthesis of a compliant component.

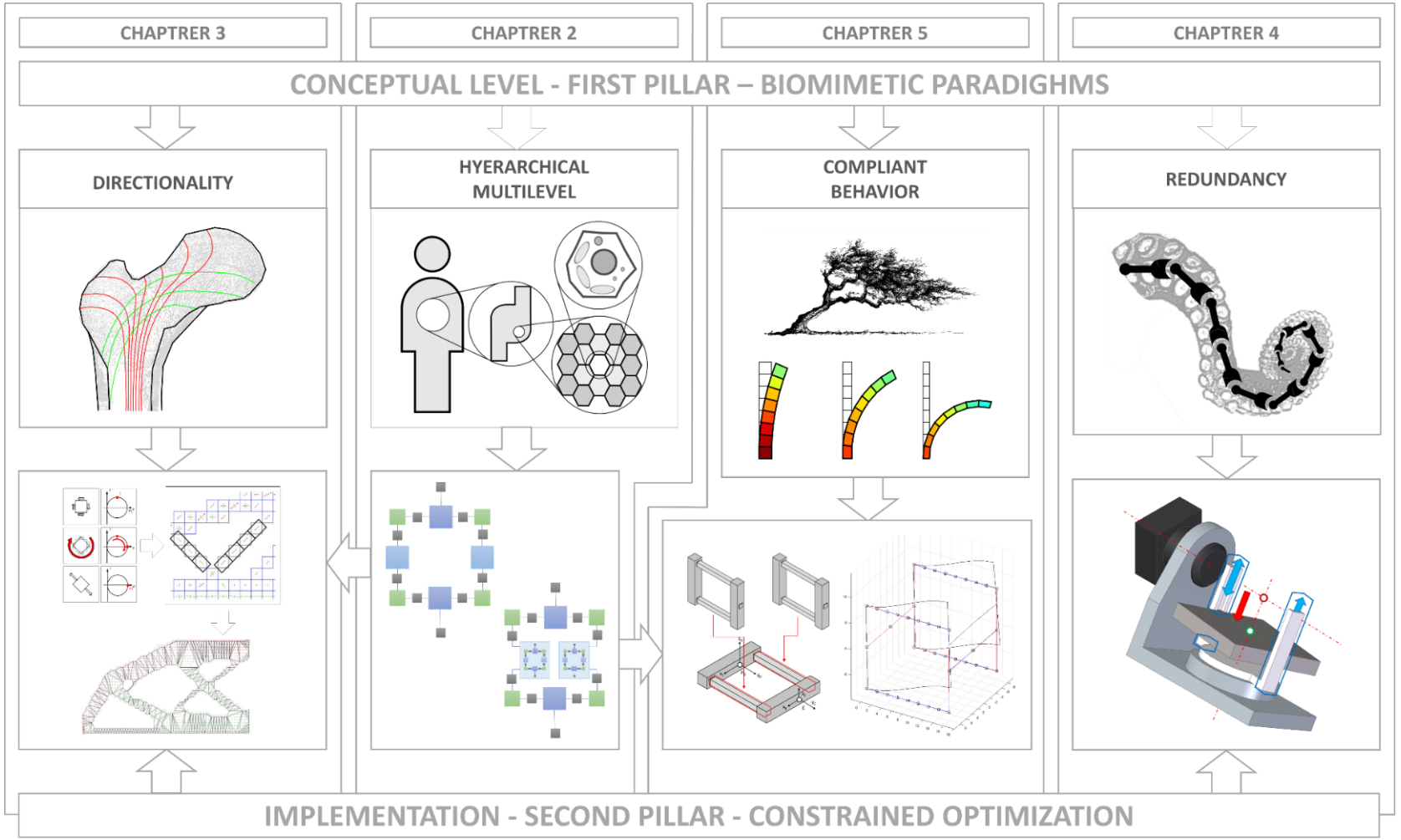


Figure 1.1: the two pillars of the research and the four biomimetic paradigms

1.4 Conclusions

The goal of the present research is providing a new methodology for the construction of tools for design. The proposal is based on the two main pillars which have been presented in this chapter: the adoption of biomimetic paradigms for technical purposes, implemented by the mean of optimization procedures.

More in detail, the main idea is taking inspiration by the living beings, in order to understand and adopt the principles they are based on for the ideation of new technical solutions. This approach is different from the one based on replicating the biological structures, which, in most of the cases, obtained limited results. On the contrary, here the attention is focused on understanding the mechanisms that allow the living being to have high performances, and implementing this knowledge in design methodology and tools. In this way, the objective is not replicate a particular tissue or limb, but overcome the limits of present technologies.

In the development of the present work, four biomimetic paradigms will be taken in account: the hierarchical multilevel organization of the biological tissues, the directionality anisotropy of the living matter, the redundancy of the limbs, the mechanical flexibility of the organic structures. Every chapter is devoted to the development of a solution to a technical problem adopting one of the previous principles, implemented by the mean of the formulation of a constrained optimization problem, which represent the mathematical tool for bridging these conceptual ideas to a practice application.

Since the objective of the research is investigating the possibility of creating bioinspired tools for the design, in the next four chapters, there have been identified four research fields suitable for the application of the four biomimetic paradigms. For each domain, it had been caried out the analysis of the corresponding states of the art, highlighting technical gaps and open issues.

As a brief resume of the contents, it can be anticipated here that chapter 2 will introduce an ontological framework which overcomes lack in the possibility of modelling biological and technical systems at different dimensional levels, typical of the existing conceptual design tools.

Chapter 3 and chapter 5 deal with the creation of original tools for the automated design of both structures and compliant mechanisms, implementing the ontological schema

disclosed in chapter 2. Finally, chapter 4 investigates the possibility of improving the design of mechanisms adopting an integrated design approach based on the concept of kinematic redundancy.

2 State of the art in Biomimetics and hierarchical multilevel approach

2.1 Introductive overview of the subject

In the past years, many engineers and researches have focused the attention on the biological systems: Biomimetics is the name of the discipline which studies the formulation of technical solutions inspired by the observation of living beings. This particular research field has been widely developed in the past years, and the reason for such interest is easily explained by the extraordinary performance of the structures and systems which may be found in nature. Biomimetics is particularly appealing under an engineering point of view because it represents the effort of taking advantage of billions of years of natural selection: during such long-time, evolution actually produced solutions which, the most of times, are superior compared to the ones produced by human technology [1].

Nevertheless, the task of artificially replicating the biological systems is not trivial. Many attempts have been done in order to copy the shape, the structure or the organization of plants and animals. Anyway, simply mimicking a particular aspect of an organism, rarely allows to obtain relevant results [2]. In this sense, it seems that the possibility of designing structures comparable to biological ones depends on the possibility of exactly replicating the complexity of the organization of the living matter: as an example, it had been highlighted that the performance of biological structures doesn’t depend on the intrinsic performance of the material involved, but by the way the material is arranged in different hierarchically organized layers [3][4].

2.2 Biomimetics in material design

Despite the difficulties, many remarkable results have been obtained implementing the biological paradigm in actual technical solutions. One of the fields, which took most from the observation of the nature, is the study of materials. For instance, many researchers investigated the mechanical performance of some specific biological structures deriving

from the relations between the various organization levels [5]. The limit of this approach is that the multilevel hierarchical organization of biological material is supported by its intrinsic capabilities of self-organizing and self-repairing, which cannot be replicated by technological materials, at the moment [6]. Nevertheless, architectured materials is a promising research field. As an example, it may be reported the study of different micro structural materials [7], which investigates the relation between the lattice topology, and its mechanical characterization, underling the evidence that hierarchical design increases buckling strength, but many other researches could be cited as well [8][9]. The majority of the works highlights that the most important thing is the characterization of materials starting from the description of their own inner structures. Starting from nano-metric dimensional scales, different properties of the materials derive from different structures, and a central topic is the relation between the different structures, and how difference parameters influence one another. This led to the creation of a new research field commonly named as computational materials [10][11]. A last remark about the design of architectured materials is that this research field has strict connections with topology optimization, which has been proposed as tool for the definition of microstructural cellular units [12][13].

2.2.1 Biomimetics in robotics

Robots are mechanisms designed to accomplish complex tasks with a certain level of autonomy. In the first applications, they have been employed for industrial purposes, in very structured environments, such as industrial lines. Later on, requirements in terms of flexibility and reliability increased, and nowadays robotic manipulators must be able to operate in complex nonstructured environments, without prescribed conditions, showing a high reactivity to external conditions changes; moreover, since interaction with humans is becoming more and more common, it is necessary to implement technical solutions in order to improve the intrinsic safety.

For these reasons, it doesn’t surprise robotics is a field in which biomimetic found a great number of applications since a long time: actually, it could be said that robotics is natively affine to biomimetic. This fact is clear if we consider the researches that, for example, try to replicate some functionalities of the human body by the means of particular

mechanisms classes [14]. Under a more general point of view, the work of Pfeifer, Iida and Bongard [15] represent a conceptual starting point for the definition of abstract categories useful for the creation of an ontology framework for the development of robotic applications: most of the principles identified in the research are evidently bio inspired, comprising the possibility of overlapping of different physical principles, or, in in other words, applying the concept of redundancy. Moreover, overcoming the classical paradigm of stiff mechanism composed by rigid links, it is possible to fulfil new functional requirements, as the intrinsic safety.

This approach is the base for the investigation of new typologies of devices: for instance, in the last years, there has been a great interest in the development of robots composed by links characterized by elastic modulus comparable to the muscles one [16]. Among the fact that these kinds of structures can fulfil different functions with to respect more rigid ones, new strategies are required in order to solve some issues related to the control, such as the management of the kinematic redundancy [17]. In fact, as it had been highlighted by the study of the solution of kinematic problems for manipulators with stiff links [18], the management of redundancy implies a rise of the control complexity level.

2.2.2 Biomimetics in conceptual design

Computational materials and bioinspired robotics are examples of the application of the biomimetic paradigm to a particular research field. Furthermore, the idea of many researchers is to generalize the possibility to formulate technical solutions inspired by nature to a wider number of engineering problems, possibly all.

Functional analysis-based frameworks are an example of the attempt of providing ontological tools to support biomimetics. In chapter 1, IDEA-INSPIRE (inspired by SAPPhIRE) [19][20], and DANE (inspired by FBS) [21] have already been reported: they represent actual applications thought to support designers and engineers to use the knowledge deriving from biology in a more systematic way. The results are remarkable, but there is still an issue related to the correspondence between the models describing biological entities, and the models used for the synthesis of human artifacts.

A suggestion in order to overcome the gap between biology and technology, may be probably found in the work of Helms [22], who states that, in order to provide a the

definition of a technical problem, it is possible adopting two approaches: the first one is the already cited functional approach, and the second is an approach based on the use of an explicit representation by the use of quantitative evaluation of the required performance, definition of objective functions, and an actual setup of an optimization problem. As already anticipated in chapter 1, the latter is one of the pillars of the present research, and, in the next section, framework will be presented that is the ontological base for a novel topology optimization procedure for the structures in chapter 3, and a novel model for the synthesis of compliant mechanisms in chapter 5.

2.3 The multi-level hierarchical organization paradigm

As introduced in chapter 1, a first important stage in order to provide effective bio inspired tools for design is the definition of a suitable ontological framework. The original proposal, disclosed in a journal paper [23], and presented in international conferences, will be described below. More in detail, for what have been highlighted in the previous section, the schema must be able to describe a hierarchical multilevel organization of the structures, which is one of the main features characterizing the living matter. In fact, in order to model a natural system, it is necessary to provide a description of the elements at all the scale levels. The dimensional levels will be conventionally called macro and micro levels. For each level there must be a core functional decomposition into structure, behaviour, and function with a particular attention to topology.

As an example, finite element analysis framework may be considered a multilevel schema with two different representation levels: in fact, the global behaviour of the main structure depends by the way the elements are arrange according to a certain topology, and by the characterization of the finite element which provides a representation of the interest field. Before describing the proposed methodology, the main bricks of the future ontology for a multilevel framework are enlisted below, and graphically resumed in figure 2.1.

2.3.1 Thesaurus of elements and relations of the hierarchical multilevel

Macro Level

As it had been already implicitly explained by introducing the method, macro level is the tier corresponding to the whole system. The first axiom of the multilevel approach states that a system can always be decomposed in a set of constitutive elements, which should not be further divided, unless considering a lower level of description. According to the external stimuli, the interaction among the elements determines the global behaviour of the system.

Micro Level

The constitutive elements of the system can be further analysed at a lower description level, the micro level indeed. In this domain, every element has its own properties (mechanical, thermal, etc…). Even if single elements are monolithic entities in the macro level representation layer, moving to the micro level, the same elements may be further decomposed, revealing their structure and their constitutive sub-elements. Furthermore, the black-box type characterization of the elements at macro level is disclosed at the micro level, since their behaviour is determinate by the mutual interaction between the sub-elements. As a last remark, the internal interactions of the elements at the macro level are considered external stimuli for the elements at a micro level.

Relation between Macro and Micro levels

According to the definition of macrolevel and microlevel, it is always possible to consider an element of the macrolevel as a system at a micro level and vice versa. This means that macro/micro level framework can be iterated, so that the lower level of the first iteration step represents the upper level for the further iteration step.

Structure

A structure is the totality of the elements that constitute the system, connected each other according to a certain topology. The two main aspects in the description of the topology are the definition of the connectivity of the elements, and the description of the position and pose of the elements in space, and in relation with the other elements.

Behaviour

The behaviour of a structure is the set of the state changes of its elements, the changes of the nature of their interactions, and changes of the structure topology. It is important to

highlight that different behaviours correspond to different external stimuli: if a system is described by a set of differential equations, different solutions correspond to different boundary and/or initial conditions.

Function

A function is the final purpose that a certain system has to fulfil at its macro level. It is important to underline that a function is not an absolute feature for the system. In fact, a more rigorous way to define a function is the following: the function of a system is a prescribed set of behaviours according to a prescribed set of external stimuli. This definition will be recalled in chapter 5 in the stiffness characterization of the compliant mechanisms.

Topology

Topology has already been introduced in the description of the concept of structure; nevertheless, it is important highlight that it is a key point in the comprehension of the behaviour of the systems. Its importance comes from the fact that two systems, composed by the same set of elements, and subject to the same stimuli, but characterized by different topologies, may have completely different behaviours, which, again, are the response of the system to the external stimuli. An example of the relevance of this fact is that different studies about auxetic materials have been developed: counting on the contribution of the particular morphologies, these microstructures are able to globally behave like materials with negative Poisson coefficient. This is the classical example of how the distribution of the material in the space allows the material itself to obtain a certain property, in opposition of its homogeneous form, and actually fulfilling a specific function.

2.3.2 Example of application

The study of the spider's orb web has been presented as a first case study of the multilevel approach: the main goal of this first research [24] was investigating the capability if the framework in explaining the superior mechanical performances of this biological material. The application of the multilevel approach in this case was the description of the structures and elements, their behaviour, and their contribution in overcoming the contradiction rising from the different functions required to the system. As main outcome of this analysis, it had been found that the functional contradiction taken in account can

be resolved by the arrangement of different elements characterized by the same organic material, but disposed according to different topologies at a lower dimensional level. The analysis details are resumed below.

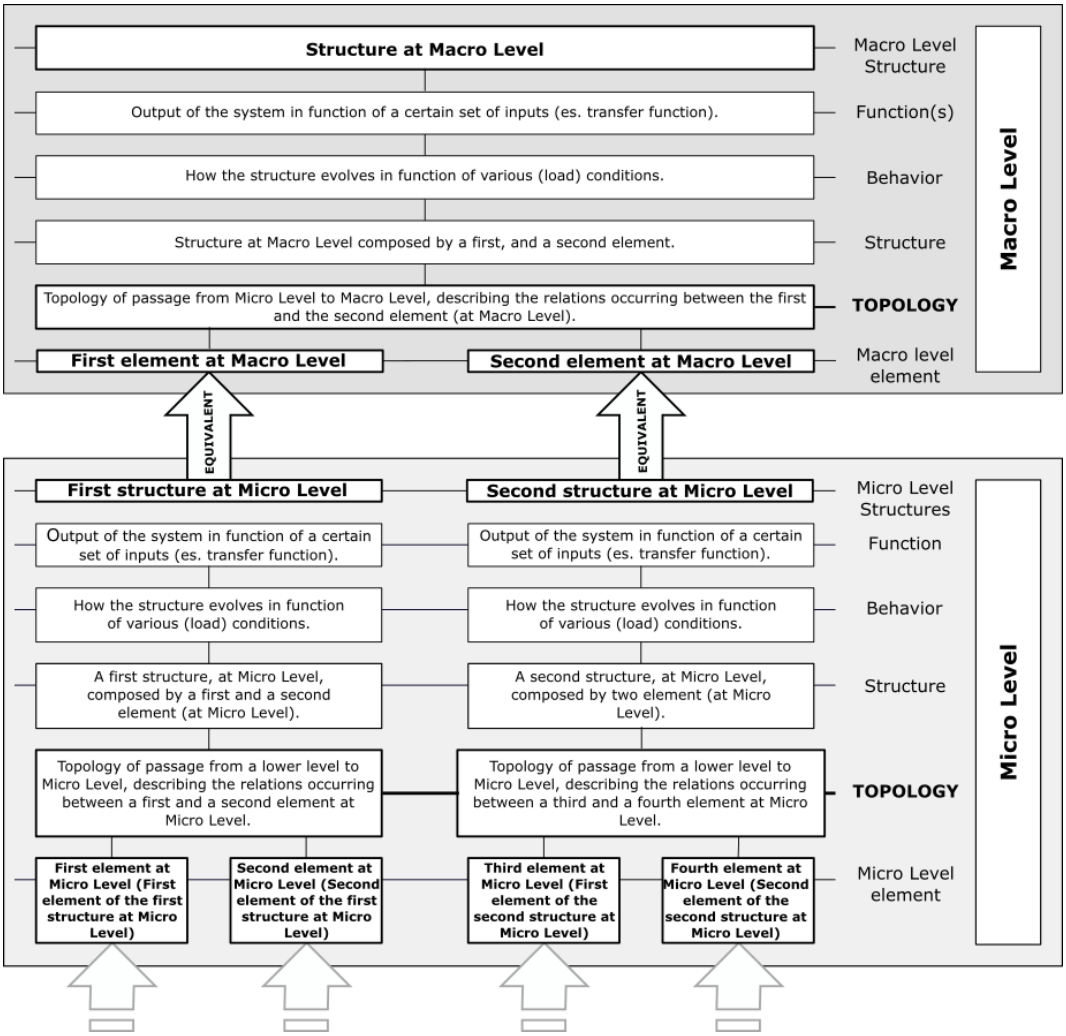


Figure 2.1: schema of the hierarchical multilevel organization of the systems

Macroscopic Level

The whole orb web is composed by two kinds of threads, having different behaviours:

- the frame silk threads are disposed in radial direction from the centre of the structure to the edge of the orb web structure. Their function is related to the preservation of the structural integrity: in fact, the maximum deformation of the dragline silk actually

corresponds to 1.25 (Ɛmax =0.25), which, according to the Clerk Maxwell's lemma, is the elongation in correspondence of the maximum force;

- the viscid silk threads, on the contrary, are arranged along a spiral path, and their main goal is ensuring the efficiency of the trap.

As a result of the particular topology adopted in the arrangement of this different elements, the orb web is able to fulfil different functions depending on different conditions:

- frame silk ensures structural integrity, saves material, provides safe lines for the spider and ensure the capability of the trap to resist the impact with the prey;
- the viscid threads are able to show a much higher elongation, and, consequently dissipate a great amount of energy.

Micro Level

As it has been stated previously, elements of a higher level are structures at a lower level: in this case, threads are the structures at the Micro Level. This means that, at the Macro Level, their role or function is providing elements characterized by different yield stresses, strength, and maximum elongations. An important remark is that frame silk and viscid silk are basically composed by the same materials produced by the gland of the animal, and their differences in terms of mechanical characteristics depends by their own topologies. In other words, the different behaviours are determined by the number and the physical properties of the elements they are made of, and the way these elements are disposed.

- frame silk is composed by a high number (4 or 5) of strands, which have high initial Young modulus, high strength, and relatively low elongation;
- viscid silk is composed of 2 strands with a lower Young modulus, a lower strength, and higher elongation. Both silks consist of elements of the same kind with a parallel disposition.

As a consequence, the mechanical response of frame and viscid silk are different, which means that, as elements, they show different responses in terms of stress-strain diagram, influencing the maximum load, and the energy dissipated, that, in both structures, is more or less equal. The viscid silk is able to absorb a certain amount of energy realizing a high elongation, and frame silk is able to absorb the same amount of energy but being less deformed.

At this level, the described structures have the function of fulfilling the requirement described at Macro Level. The mechanical responses of the frame and viscid threads, which are structures at the Micro Level, obviously correspond to the characteristics of the components at the macro level. In particular, for the viscid silk, the higher admissible force is in correspondence of a very high elongation, while for frame silk, the higher admissible force is in correspondence of a compatible elongation around 1,25.

Concluding, the adoption of two different topologies at micro level for the result of the arrangement of the basic material of the silk allows the creation of a structure able to fulfil two different functions at the macro level.

In the next chapter, the strategy of differentiating the topological arrangement of the elements at a low dimensional level will be applied to structures design; it is well known that the better performances for structural purposes are obtained realizing objects, which, at a microscopic level, are constituted by elements (or cells), which are not isotropic, so that the singular elementary unit may globally have, along the three principal directions, different values of Young and Poisson moduli. Furthermore, the orientation of the single constitutive elements may lie along the principal direction for the stress tensor field, which has been obtained by the macroscopic analysis.

Figure 2.2 provides a summary of the previous observations, placing them in the diagram shown in the figure 2.1.

The orb web is a biomimetic example that can be an inspiration for the conception of a new design methodology. According to the multilevel philosophy, the main idea is involving different optimization strategies at different dimensional levels, in order to take into account the evolutionary trend of the systems, which imposes the reordering of material at different hierarchical levels. For this reason, the goal of the next chapter will be the study of a multilevel topology optimization tool.

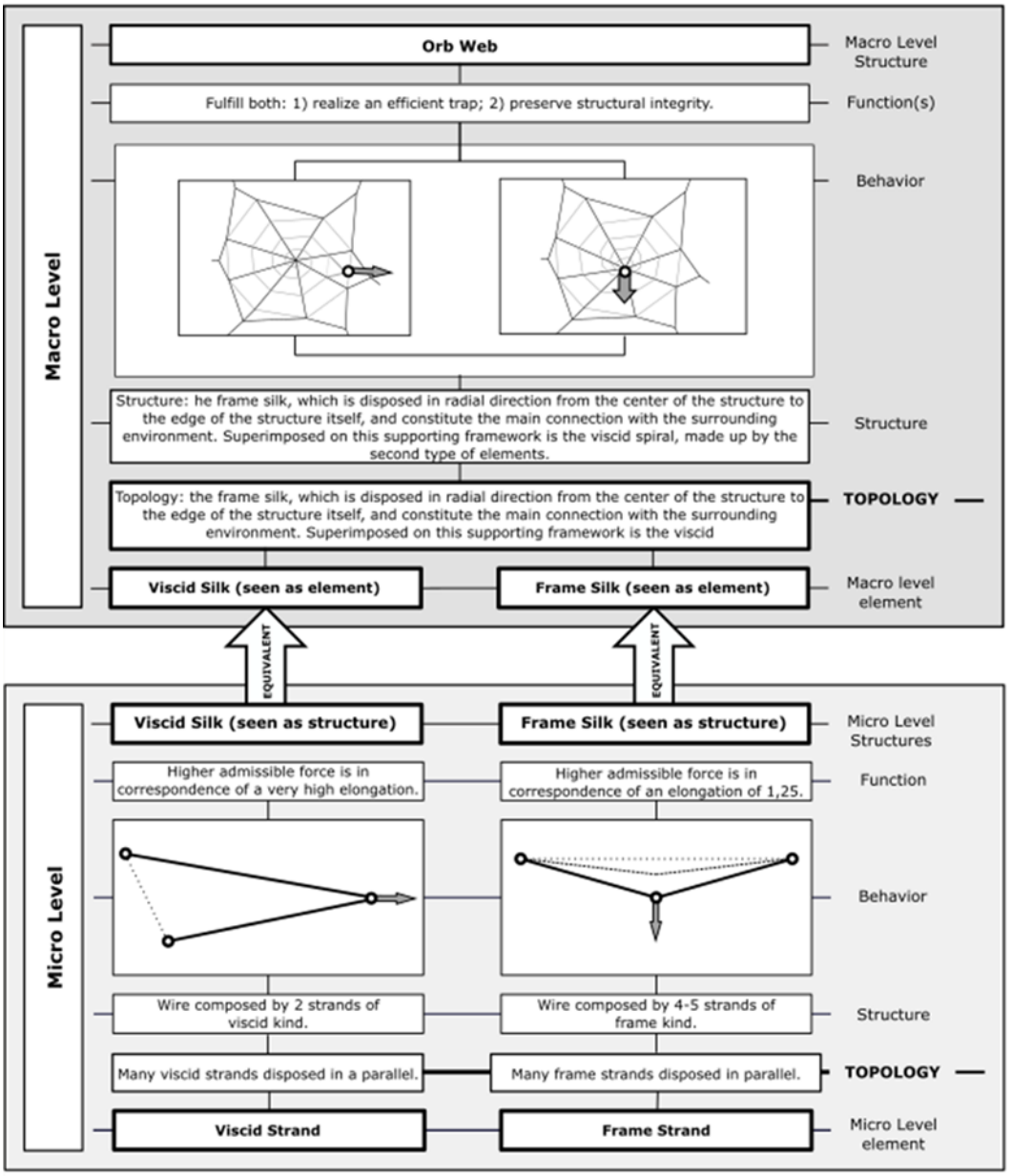


Figure 2.2: example of the hierarchical multilevel approach applied to the analysis of the spider orb web

2.4 Conclusions and preface to chapter 3

Chapter 2 has been devoted to the investigation of a possible implementation of a design approach inspired by the biomimetic hierarchical organization of living beings. Since the multilevel paradigm may be observed in the structures and functions of almost all biological systems, conceptual design seems to be the most appropriate investigation field for its application. Actually, the literature analysis regarding conceptual design tools

revealed that applications adopting a classical functional analysis-based setup, such as IDEA-INSPIRE and DANE, are not specifically designed in order to model systems organized in many hierarchical levels, under both functional and structural point of view. Consequently, the proposed methodology represents a substantial improvement with respect to the state of the art.

The first part of the present chapter reported a general brief overview of biomimetic researches; moreover, it had been provided a focus on the concept of multilevel, and its use in different research fields. It had been noticed that the hierarchical organization of the biological structures is a key factor for the determination of their performances. Furthermore, it had been highlighted that deducing technical solutions inspired by organic multilevel architectures is not that easy without a solid methodological framework.

For this reason, the second part of the chapter has been devoted to the definition of a first basic ontology, characterized by a set of elements and relations for the description of physical models at different dimensional levels. The proposed framework has been developed according to the already existing mono-level FBS framework, which has been integrated with new conceptual entities: the most important is the topology, which actual provides a connection between two different dimensional layers.

An exemplification of the multilevel approach has been given by the analysis of a biological structure. The example shows how it is possible to have a structure which is able to fulfil different contrasting requirements by increase the number of hierarchical levels of organization: this is done by the development of different topologies for the arrangement of the material at different dimension, avoiding the necessity of using many different materials.The future application of this approach is the implementation of the ontological framework in new tools for analysis and design. In the next chapter, the actual possibility of the application of the multilevel paradigm will be investigated, and this will be done by introducing an original topology optimization procedure.

2.5 Bibliography

[1] Bar-Cohen, Y. (2006). Biomimetics—using nature to inspire human innovation. Bioinspiration & biomimetics, 1(1), P1.

[2] Vincent, J. F., Bogatyreva, O. A., Bogatyrev, N. R., Bowyer, A., & Pahl, A. K. (2006). Biomimetics: its practice and theory. Journal of the Royal Society Interface, 3(9), 471-482.

[3] Meyers, M. A., Chen, P. Y., Lin, A. Y. M., & Seki, Y. (2008). Biological materials: structure and mechanical properties. Progress in Materials Science, 53(1), 1-206.

[4] Bhushan, B. (2009). Biomimetics: lessons from nature–an overview.

[5] Wegst, U. G. K., & Ashby, M. F. (2004). The mechanical efficiency of natural materials. Philosophical Magazine, 84(21), 2167-2186.

[6] Raabe, D., Sachs, C., & Romano, P. J. A. M. (2005). The crustacean exoskeleton as an example of a structurally and mechanically graded biological nanocomposite material. Acta Materialia, 53(15), 4281-4292.

[7] Fleck, N. A., Deshpande, V. S., & Ashby, M. F. (2010). Micro-architectured materials: past, present and future. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 466(2121), 2495-2516.

[8] Fish, J., & Shek, K. (2000). Multiscale analysis of composite materials and structures. Composites Science and Technology, 60(12-13), 2547-2556.

[9] Fish, J., & Wagiman, A. (1993). Multiscale finite element method for a locally nonperiodic heterogeneous medium. Computational mechanics, 12(3), 164-180.

[10] Gates, T. S., Odegard, G. M., Frankland, S. J. V., & Clancy, T. C. (2005). Computational materials: multi-scale modeling and simulation of nanostructured materials. Composites Science and Technology, 65(15-16), 2416-2434.

[11] Hafner, J. (2000). Atomic-scale computational materials science. Acta Materialia, 48(1), 71-92.

[12] Huang, X., Zhou, S. W., Xie, Y. M., & Li, Q. (2013). Topology optimization of microstructures of cellular materials and composites for macrostructures. Computational Materials Science, 67, 397-407.

[13] Sivapuram, R., Dunning, P. D., & Kim, H. A. (2016). Simultaneous material and structural optimization by multiscale topology optimization. Structural and multidisciplinary optimization, 54(5), 1267-1281.

[14] Di Gregorio, R., Parenti-Castelli, V., O'Connor, J. J., & Leardini, A. (2007). Mathematical models of passive motion at the human ankle joint by equivalent

spatial parallel mechanisms. Medical & biological engineering & computing, 45(3), 305-313.

[15] Pfeifer, R., Iida, F., & Bongard, J. (2005). New robotics: Design principles for intelligent systems. Artificial life, 11(1-2), 99-120.

[16] Kim, S., Laschi, C., & Trimmer, B. (2013). Soft robotics: a bioinspired evolution in robotics. Trends in biotechnology, 31(5), 287-294.

[17] Walker, I. D., Dawson, D. M., Flash, T., Grasso, F. W., Hanlon, R. T., Hochner, B., ... & Zhang, Q. M. (2005, May). Continuum robot arms inspired by cephalopods. In Unmanned Ground Vehicle Technology VII (Vol. 5804, pp. 303-314). International Society for Optics and Photonics.

[18] Lenarčič, J. (2014, September). Some computational aspects of robot kinematic redundancy. In International Conference on Parallel Problem Solving from Nature (pp. 1-10). Springer, Cham.

[19] Chakrabarti, A., Sarkar, P., Leelavathamma, B., & Nataraju, B. S. (2005). A functional representation for aiding biomimetic and artificial inspiration of new ideas. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 19(2), 113-132.

[20] Sartori, J., Pal, U., & Chakrabarti, A. (2010). A methodology for supporting "transfer" in biomimetic design. AI EDAM, 24(4), 483-506.

[21] Vattam, S., Wiltgen, B., Helms, M., Goel, A. K., & Yen, J. (2011). DANE: fostering creativity in and through biologically inspired design. In Design Creativity 2010 (pp. 115-122). Springer, London.

[22] Helms, M., Vattam, S. S., & Goel, A. K. (2009). Biologically inspired design: process and products. Design studies, 30(5), 606-622.

[23] Caputi, A., Russo, D., & Rizzi, C. (2018). Multilevel topology optimization. Computer-Aided Design and Applications, 15(2), 193-202.

[24] Russo, D., & Caputi, A. (2016). Multilevel triz contradiction in biomimetics. In The 7th International Conference on Systematic Innovation (ICSI 2016) (pp. 490-501). Universidad NOVA de Lisboa.

3 Creation of a design tool based on biomimetic: the topology optimization revised

In this chapter the theoretical background and the implementation of the directionality paradigm inspired by biomimetic studies is described. As it had been previously stated, the main goal of the present research is delivering tools and methodologies for the design, and topology optimization is one of the most promising fields related to the automatic generation of structures, which, nowadays, is integrated in the most of FEA packages.

Such wide diffusion depends by the fact that, in its most common formulation, which implies the compliance minimization, the implementation of topology optimization in a standard finite element code is trivial. Nevertheless, minimization of the work done by the external forces is not the only requirement to take in account while designing a structure: in engineering practice is important, for instance, ensuring the safety requirements taking in account a resistance criterion for the material, or providing designs which may be manufactured with available technologies.

Next sections will present an overview of well-established techniques, and less known approaches used in the optimization of the structures design, trying to highlight the common theoretical aspects, in order to provide the most general schema in which the directionality idea will be integrated.

3.1 Theoretical setup of topology optimization

Structural optimization aims to define the ideal design of structures under prescribed boundary conditions. At a first glance, structural optimization may be divided in size optimization, shape optimization and topology optimization [25].

Size optimization consists in defining the dimensional parameters of the elements connected, according with a given topology, or, in other words, a given connectivity of the elements of the structure. Historically, the study of ground structures is a typical problem of size optimization. Shape optimization deals with the definition of the shape of the boundary of the structure, without the possibility to create or eliminate holes. Finally, topology optimization aims to define the ideal distribution of material in a given

volume of space, without limitations, except the indication of a design space, loads, and mechanical constraints.

Such definition of topology optimization is general, provided at a very high abstraction level, and, practically, is not very useful for the implementation of an automatic generation tool. In order to set up a first theoretical framework, a specific optimization problem is presented, which will be the starting point for the further reasoning. The specific problem is the minimum compliance design.

Let us consider a structure which occupies a region $\Omega \in \mathbb{R}^2 \; or \; \mathbb{R}^3$, and which is subject to certain constraints and applied forces. The goal is determining the ideal stiffness tensor E_{ijhk} in all the design domain Ω. For this purpose, a linear form is introduced

$$W(\underline{x}) = \int_{\Omega} \underline{x}^T \cdot \underline{b} \, d\Omega + \int_{\Gamma} \underline{x}^T \cdot \underline{t} \, d\Gamma \tag{3.1}$$

with $\underline{b}$ denoting the body forces field, and $\underline{t}$ denoting the forces acting at the boundary, and a bilinear form:

$$U\left(\underline{y}, \underline{z}\right) = \int_{\Omega} \varepsilon_{ij}\left(\underline{y}\right) E_{ijhk} \varepsilon_{hk}(\underline{z}) d\Omega \tag{3.2}$$

with

$$\varepsilon_{ij}(\underline{x}) = \underline{\varepsilon}(\underline{x}) = \frac{1}{2}\left(\underline{\nabla} \cdot \underline{x} + (\underline{\nabla} \cdot \underline{x})^T\right) \tag{3.3}$$

If the vectors $\underline{v}$ and $\underline{w}$ are such that $\underline{v}, \underline{w} \in U_{adm}$ where U_{adm} is the set of kinematically admissible displacements, then the equation

$$U(\underline{v}, \underline{w}) = W(\underline{w}) \tag{3.4}$$

is the weak formulation, or the variational formulation, of the equilibrium of the deformable continuum. It may be recalled here that a weak form is a notation which involves the use of an integral expression [26], and it represents the evaluation of the average value of the physical quantity, instead of providing the pointwise value.

Under such conditions, the minimum compliance optimization problem can be formulated as follows:

$$\begin{aligned} min_{\underline{v} \in U_{adm}, \underline{\underline{E}}} \quad & W(\underline{w}) \\ s.t. \quad & U(\underline{v}, \underline{w}) = W(\underline{w}) \quad \forall \underline{v} \in U_{adm} \\ & \underline{\underline{E}} \in \underline{\underline{E}}_{adm} \end{aligned} \tag{3.5}$$

where the relation $\underline{\underline{E}} \in \underline{\underline{E}}_{adm}$ denotes the fact that the tensor is such that it is able to describe the characterization of the given material. It is important to address that, in this formulation, there are two fields of interest: the first one is the tensor field $\underline{\underline{E}}$, and the second one is the displacement field $\underline{v}$. In order to solve the given problem by computational means, the most common approach is the use of the finite elements, and the most important remark about this is that, if the same element mesh is used for both $\underline{\underline{E}}$ and $\underline{v}$ fields, the implementation of the solution results to be straight forward. Actually, this is the reason of the diffuse implementation of topology optimization in finite element software. Furthermore, such easy implementation is due to the fact that the optimization problem involves the compliance as function to minimize: as we will see later on, this particular problem is easy to solve since it is a so called self-adjoint problem. Other formulations are much more challenging, such as stress-based optimization or design of compliant mechanisms.

3.2 Finite element method fundament

As stated in the previous section, the finite element method is involved in the actual solution of topology optimization problems; the procedure is based on a weak formulation of the principle of stationarity of the potential energy of a system. The potential energy can be expressed as the sum of the internal potential energy U_i and the external potential energy U_e:

$$U = U_i + U_e \tag{3.6}$$

Actually, in this dissertation, the internal potential energy is the elastic energy stored in the structure itself, and the external potential energy is minus the potential energy of the forces applied to the structure:

$$U_e = -W \tag{3.7}$$

A configuration of the system is an equilibrium configuration if the vector of the lagrangian variables $\underline{x}$, describing the state of the system, is a stationary point for the potential energy:

$$\delta U(\underline{x}) = \delta U_i(\underline{x}) - \delta W(\underline{x}) = 0 \tag{3.8}$$

In general, the state of a structure can be described by the displacement field $\underline{u}$ generated by the application of external forces, and imposing prescribed nodal displacements in correspondence of kinematic constraints. In this framework, the Lagrange variables are the displacements. This is a very general principle, and, in order to use it for computational purposes, it is possible to use the Rayleigh-Ritz formulation, which is based on the use of the shape functions for interpolating the interest fields, i.e., the nodal displacements $\underline{v}$, for the analysis. It may be recalled here that an alternative approach is the one based on the particular formulation of the weighted residuals method proposed by Galerkin, which provides results similar to the Rayleigh-Ritz approach: actually, the stiffness matrix results the same.

In the finite element method, one of the crucial stages is the definition of the type of the discretization elements. The development of a general treatise is far beyond the purposes of this work, and here will be provided the development of the fundamental equation for the case of the bidimensional isoperimetric four nodes element, which is the basic element for the development of the meshes that will be used later on.

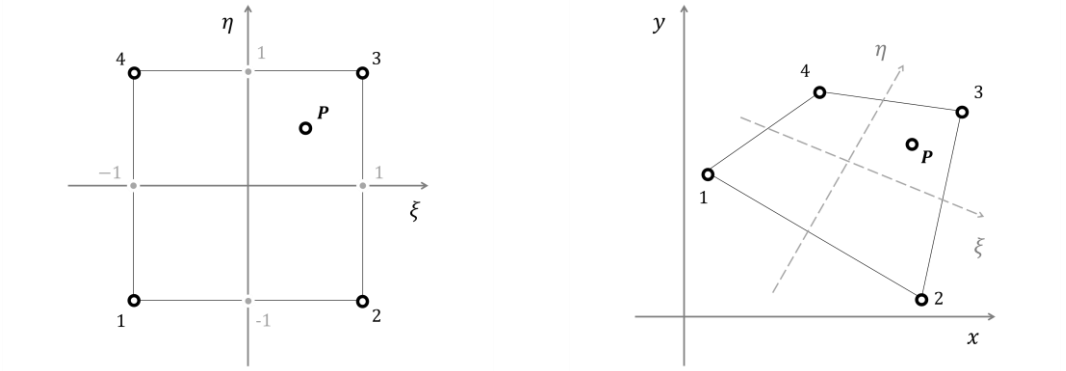
The shape functions can be expressed by the following [27]:

$$N_i(\xi,\eta) = \frac{1}{4}(1+\xi_i\xi)\cdot(1+\eta_i\eta) \quad i = 1\,,2\,,3\,,4 \tag{3.9}$$

With shape functions, it is possible to interpolate any field using the values of the field itself in correspondence of the four nodes of the element:

$$f(\xi,\eta) = \sum_{i=1}^{4} f_i\, N_i(\xi,\eta) \tag{3.10}$$

and it is done by defining a mapping between the so-called master element, and the slave element, which is the actual component of the mesh. The correspondence between the points of the master element and the slave element are depicted in figure 3-1 (a) and (b) respectively:



(a) *(b)*

Figure 3.1: master (a) and slave element (b) in isoparametric mapping

If we consider the coordinate x and the coordinate y of the points belonging the slave element as two different fields, it is possible to write:

$$\begin{aligned} x(\xi,\eta) &= x_1 \cdot N_1(\xi,\eta) + x_2 \cdot N_2(\xi,\eta) + x_3 \cdot N_3(\xi,\eta) + x_4 \cdot N_4(\xi,\eta) \\ y(\xi,\eta) &= y_1 \cdot N_1(\xi,\eta) + y_2 \cdot N_2(\xi,\eta) + y_3 \cdot N_3(\xi,\eta) + y_4 \cdot N_4(\xi,\eta) \end{aligned} \tag{3.11}$$

This is a mapping from the space of the master element to the space of the slave element, which is defined by functions of class C^1. This means that it is possible to write the Jacobian of the transformation:

$$\underline{\underline{J}}(\xi,\eta) = \begin{vmatrix} \frac{\partial x}{\partial \xi}(\xi,\eta) & \frac{\partial y}{\partial \xi}(\xi,\eta) \\ \frac{\partial x}{\partial \eta}(\xi,\eta) & \frac{\partial y}{\partial \eta}(\xi,\eta) \end{vmatrix} = \begin{vmatrix} \sum_{i=1}^{4} x_i \frac{\partial N_i(\xi,\eta)}{\partial \xi}(\xi,\eta) & \sum_{i=1}^{4} y_i \frac{\partial N_i(\xi,\eta)}{\partial \xi}(\xi,\eta) \\ \sum_{i=1}^{4} x_i \frac{\partial N_i(\xi,\eta)}{\partial \eta}(\xi,\eta) & \sum_{i=1}^{4} y_i \frac{\partial N_i(\xi,\eta)}{\partial \eta}(\xi,\eta) \end{vmatrix} \tag{3.12}$$

On the other hand, it is possible to express the two components of the displacement of the generic point of the element in the same way:

$$\begin{aligned} u(\xi,\eta) &= u_1 \cdot N_1(\xi,\eta) + u_2 \cdot N_2(\xi,\eta) + u_3 \cdot N_3(\xi,\eta) + u_4 \cdot N_4(\xi,\eta) \\ v(\xi,\eta) &= v_1 \cdot N_1(\xi,\eta) + v_2 \cdot N_2(\xi,\eta) + v_3 \cdot N_3(\xi,\eta) + v_4 \cdot N_4(\xi,\eta) \end{aligned} \tag{3.13}$$

Resuming, the use of the shape functions allows to express the position vector of the generic point $\underline{x} = (x,y)^T = \big(x(\xi,\eta), y(\xi,\eta)\big)^T$ and its generic displacement $\underline{v} = (u,v)^T = \big(u(\xi,\eta), v(\xi,\eta)\big)^T$ in functions of their values in the nodes of the element:

$$\underline{x}(\xi,\eta) = \begin{vmatrix} x(\xi,\eta) \\ y(\xi,\eta) \end{vmatrix} =$$

$$= \begin{vmatrix} N_1(\xi,\eta) & 0 & N_2(\xi,\eta) & 0 & N_3(\xi,\eta) & 0 & N_4(\xi,\eta) & 0 \\ 0 & N_1(\xi,\eta) & 0 & N_{2(\xi,\eta)} & 0 & N_3(\xi,\eta) & 0 & N_4(\xi,\eta) \end{vmatrix} \cdot \begin{vmatrix} x_1 \\ y_1 \\ \vdots \\ x_4 \\ y_4 \end{vmatrix} \tag{3.14}$$

and

$$\underline{v}(\xi,\eta) = \begin{vmatrix} u(\xi,\eta) \\ v(\xi,\eta) \end{vmatrix} =$$

$$= \begin{vmatrix} N_1(\xi,\eta) & 0 & N_2(\xi,\eta) & 0 & N_3(\xi,\eta) & 0 & N_4(\xi,\eta) & 0 \\ 0 & N_1(\xi,\eta) & 0 & N_{2(\xi,\eta)} & 0 & N_3(\xi,\eta) & 0 & N_4(\xi,\eta) \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ v_1 \\ \vdots \\ u_4 \\ v_4 \end{vmatrix} \tag{3.15}$$

$$= \underline{\underline{N}} \cdot \underline{u_E}$$

This formulation is very useful because it allows to express all the necessary mechanical quantities in function of the nodal displacements. In fact, recalling the definition of strain, it is possible to write

$$\left\{ \varepsilon_{xx} = \frac{\partial u}{\partial x} \; ; \; \varepsilon_{yy} = \frac{\partial v}{\partial y} \; ; \; \gamma_{xx} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right. \implies \begin{vmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \gamma_{xy} \end{vmatrix} = \begin{vmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{vmatrix} \cdot \begin{vmatrix} u \\ v \end{vmatrix} = \underline{\underline{L}} \cdot \underline{v} \tag{3.16}$$

keeping in mind that the compatibility equation must be fulfilled:

$$\frac{\partial^2 \varepsilon_{xx}}{\partial x^2} + \frac{\partial^2 \varepsilon_{yy}}{\partial y^2} = 2 \cdot \frac{\partial^2 \gamma_{xy}}{\partial x \partial y} \tag{3.17}$$

Recalling the definition of the shape functions, it is possible to define the $\underline{\underline{B}}$ matrix:

$$\begin{vmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \gamma_{xy}\end{vmatrix} = \begin{vmatrix}\frac{\partial}{\partial x} & 0\\ 0 & \frac{\partial}{\partial y}\\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x}\end{vmatrix} \cdot \begin{vmatrix}N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4 & 0\\ 0 & N_1 & 0 & N_2 & 0 & N_3 & 0 & N_4\end{vmatrix} \cdot \begin{vmatrix}u_1\\ v_1\\ \vdots\\ u_4\\ v_4\end{vmatrix} \Rightarrow \tag{3.18}$$

$$\Rightarrow \quad \underline{\underline{B}} = \underline{\underline{L}} \cdot \underline{\underline{N}}$$

The matrix $\underline{\underline{B}}$ provides a relation between the strain in correspondence of a certain point of the master element space, and the nodal displacements:

$$\underline{\underline{B}} = \begin{vmatrix}\frac{\partial N_1}{\partial x} & 0 & \frac{\partial N_2}{\partial x} & 0 & \frac{\partial N_3}{\partial x} & 0 & \frac{\partial N_4}{\partial x} & 0\\ 0 & \frac{\partial N_1}{\partial y} & 0 & \frac{\partial N_2}{\partial y} & 0 & \frac{\partial N_3}{\partial y} & 0 & \frac{\partial N_4}{\partial y}\\ \frac{\partial N_1}{\partial y} & \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial y} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial y} & \frac{\partial N_3}{\partial x} & \frac{\partial N_4}{\partial y} & \frac{\partial N_4}{\partial x}\end{vmatrix} \Rightarrow \tag{3.19}$$

$$\Rightarrow \underline{\varepsilon} = \begin{vmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \gamma_{xy}\end{vmatrix} = \underline{\underline{B}} \cdot \underline{u_E}$$

Furthermore, it is possible to express an infinitesimal variation of the strain tensor, and an infinitesimal variation of the displacement vector of a generic point in function of an infinitesimal variation of the vector of the nodal displacements:

$$\begin{aligned}\delta\underline{\varepsilon} &= \delta\left(\underline{\underline{B}} \cdot \underline{u}\right) = \underline{\underline{B}} \cdot \delta\underline{u_E}\\ \delta\underline{v} &= \delta\left(\underline{\underline{N}} \cdot \underline{u}\right) = \underline{\underline{N}} \cdot \delta\underline{u_E}\end{aligned} \tag{3.20}$$

(notice that $\delta\underline{\varepsilon} = \delta\underline{\varepsilon}(\xi,\eta)$ and $\delta\underline{v} = \delta\underline{v}(\xi,\eta)$ recalling that $\underline{\underline{N}} = \underline{\underline{N}}(\xi,\eta)$ and $\underline{\underline{B}} = \underline{\underline{B}}(\xi,\eta)$).

Moreover, for the linear elastic problem, the stress/strain relation is expressed as follows:

$$\underline{\sigma} = \begin{vmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \tau_{xy}\end{vmatrix} = \frac{E}{1-v^2} \cdot \begin{vmatrix}1 & v & 0\\ v & 1 & 0\\ 0 & 0 & \frac{1-v}{2}\end{vmatrix} \cdot \begin{vmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \gamma_{xy}\end{vmatrix} = \underline{\underline{D}} \cdot \underline{\varepsilon} \tag{3.21}$$

Using the principle of the virtual works, where the force system is provided by the forces applied to the structure, and the displacement system is provided by the tensor field $\delta\underline{\varepsilon}$ and the vector field $\delta\underline{v}$, it is possible to write:

$$\int_{\Omega_E} (\delta\underline{\varepsilon})^T \cdot \underline{\sigma}\, d\Omega = \int_{\Omega_E} (\delta\underline{v})^T \cdot \underline{b}\, d\Omega + \int_{\Gamma} (\delta\underline{v})^T \cdot \underline{t}\, d\Gamma + \sum_{i=1}^{4} \delta\underline{v}\big|_{\underline{x}=\underline{x}_i} \cdot f_i \tag{3.22}$$

with $\underline{b}$ denoting the body forces field, and $\underline{t}$ denoting the forces acting at the boundary, both null in this case, and f_i denoting the concentrated nodal forces.

Under this hypothesis, and recalling the expressions for $\delta\underline{\varepsilon}$ and $\delta\underline{v}$, the equilibrium of the structure reads as:

$$\begin{aligned} &\int_{\Omega_E} (\delta\underline{\varepsilon})^T \cdot \underline{\sigma}\, d\Omega = \int_{\Omega_E} (\delta\underline{\varepsilon})^T \cdot \underline{\underline{D}} \cdot \underline{\varepsilon}\, d\Omega = \int_{\Omega_E} (\delta\underline{u}_e)^T \cdot \left(\underline{\underline{B}}\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}} \cdot \underline{u}_e\, d\Omega = \\ &= \sum_{i=1}^{N} \delta\underline{v}\big|_{\underline{x}=\underline{x}_i} \cdot f_i = (\delta\underline{u}_e)^T \cdot \underline{f}_e \end{aligned} \tag{3.23}$$

By definition, the components of the vectors $\underline{u}_E$ and $\delta\underline{u}_E$ are not dependent by the point coordinate inside the element, and then, finally the following holds:

$$\begin{aligned} &(\delta\underline{u}_e)^T \cdot \left[\int_{\Omega} \left(\underline{\underline{B}}\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}\, d\Omega\right] \cdot \underline{u}_e = (\delta\underline{u}_e)^T \cdot \underline{f} \quad \Longrightarrow \quad \underline{\underline{K}} \cdot \underline{u}_e = \underline{f}_e \\ &\underline{\underline{K}}_e = \int_{\Omega_E} \left(\underline{\underline{B}}\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}\, d\Omega \end{aligned} \tag{3.24}$$

The matrix $\underline{\underline{K}}_E$ is the element stiffness matrix, and, in other to carry out its computation, it is necessary to calculate the integral in the coordinate space (ξ, η) of the master element. By definition, matrix $\underline{\underline{B}}$ is:

$$\underline{\underline{B}} = \underline{\underline{B}}(\xi,\eta) = \begin{vmatrix} \frac{\partial N_1}{\partial x}(\xi,\eta) & 0 & \cdots & \frac{\partial N_4}{\partial x}(\xi,\eta) & 0 \\ 0 & \frac{\partial N_1}{\partial y}(\xi,\eta) & \cdots & 0 & \frac{\partial N_4}{\partial y}(\xi,\eta) \\ \frac{\partial N_1}{\partial y}(\xi,\eta) & \frac{\partial N_1}{\partial x}(\xi,\eta) & \cdots & \frac{\partial N_4}{\partial y}(\xi,\eta) & \frac{\partial N_4}{\partial x}(\xi,\eta) \end{vmatrix} \tag{3.25}$$

which means that it is necessary compute the derivative of the shape functions with to respect the coordinates of the slave element. To do this, it is necessary derive the shape functions with to respect the coordinates of the master element, and then apply the derivative chain rule:

$$\frac{\partial N_i}{\partial \xi} = \frac{\partial N_i}{\partial x}\cdot\frac{\partial x}{\partial \xi} + \frac{\partial N_i}{\partial y}\cdot\frac{\partial y}{\partial \xi} \quad ; \quad \frac{\partial N_i}{\partial \eta} = \frac{\partial N_i}{\partial x}\cdot\frac{\partial x}{\partial \eta} + \frac{\partial N_i}{\partial y}\cdot\frac{\partial y}{\partial \eta} \quad \Longrightarrow$$

$$\begin{vmatrix} \frac{\partial N_i}{\partial \xi} \\ \frac{\partial N_i}{\partial \eta} \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial \xi}(\xi,\eta) & \frac{\partial y}{\partial \xi}(\xi,\eta) \\ \frac{\partial x}{\partial \eta}(\xi,\eta) & \frac{\partial y}{\partial \eta}(\xi,\eta) \end{vmatrix} \cdot \begin{vmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \end{vmatrix} = \underline{\underline{J}}(\xi,\eta) \cdot \begin{vmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \end{vmatrix} \Longrightarrow \tag{3.26}$$

$$\begin{vmatrix} \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial \xi}(\xi,\eta) & \frac{\partial y}{\partial \xi}(\xi,\eta) \\ \frac{\partial x}{\partial \eta}(\xi,\eta) & \frac{\partial y}{\partial \eta}(\xi,\eta) \end{vmatrix}^{-1} \cdot \begin{vmatrix} \frac{\partial N_i}{\partial \xi} \\ \frac{\partial N_i}{\partial \eta} \end{vmatrix} = \left(\underline{\underline{J}}(\xi,\eta)\right)^{-1} \cdot \begin{vmatrix} \frac{\partial N_i}{\partial \xi} \\ \frac{\partial N_i}{\partial \eta} \end{vmatrix}$$

Once the derivatives of the shape functions N_i with respect to ξ and η are computed, it is possible to find a new expression for the element stiffness matrix:

$$\begin{aligned} \underline{\underline{K_e}} &= \int_{\Omega_E} \left(\underline{\underline{B}}\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}\, d\Omega = \int_{\Omega_E} \left(\underline{\underline{B}}(\xi,\eta)\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}(\xi,\eta) \cdot t\, d\mathrm{xdy} = \\ &= t \int_{-1}^{1}\int_{-1}^{1} \left(\left(\underline{\underline{B}}(\xi,\eta)\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}(\xi,\eta) \cdot det\left(\underline{\underline{J}}(\xi,\eta)\right)\right) d\xi \mathrm{d}\eta \end{aligned} \tag{3.27}$$

where t is the thickness of the plate, so that the volume of the element is:

$$\int_{\Omega_E} t\, d\mathrm{xdy} = t \int_{-1}^{1}\int_{-1}^{1} det\left(\underline{\underline{J}}(\xi,\eta)\right) d\xi \mathrm{d}\eta \tag{3.28}$$

In the finite element method, in order to efficiently compute the integral, it is common practice the use of the integration by Gauss points. Since the description of the theoretical basis of the numerical integration procedure is far beyond the purpose of this section, here is reported the final result, which is the expression of the element rigidity matrix is:

$$\underline{\underline{K_E}} = t \sum_i \sum_j w_i w_j \left(\underline{\underline{B}}(\xi_i,\eta_j)\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}(\xi_i,\eta_j) \cdot det\left(\underline{\underline{J}}(\xi,\eta)\right) \tag{3.29}$$

where the parameters w_i, w_j, ξ_i, η_j are pre-calculated values, depending on the number of Gauss points involved in the computation. $\underline{\underline{K_E}}$ is the element stiffness matrix, which has 8x8 dimension for the 4 node isoparametric element.

In order to solve the whole structure, the element stiffness matrixes must be assembled in a global stiffness matrix, which has MxM dimension: M is the number of degrees of freedom of the structure, which do not have any prescribed displacement; in other words, in the case of isoparametric bilinear elements, M is equal to 2 times the number of the nodes minus the number of bounded degrees of freedom. The global stiffness matrix may be expressed as:

$$\underline{\underline{K}} = \sum_i \underline{\underline{K'_{e(i)}}} \tag{3.30}$$

where the matrix $\underline{\underline{K'_{E(i)}}}$ is the i-th element stiffness matrix at the global level (dimension MxM).

Once the global stiffness matrix has been assembled, it is possible to solve the structure, determining the vector of all the nodal displacements $\underline{u}$ (vector of dimension N), knowing the boundary conditions and the global vector of the external forces $\underline{f}$:

$$\underline{\underline{K}} \cdot \underline{u} = \underline{f} \quad \Longrightarrow \quad \underline{u} = \left(\underline{\underline{K}}\right)^{-1} \cdot \underline{f} \tag{3.31}$$

where $\underline{u}$ and $\underline{f}$ are vectors of dimension Mx1.

3.3 Relaxation of the binary optimization problem and material interpolation.

Let us recall now that the topology optimization problem is the determination of the ideal material layout within a feasible design space. Ideally, this is a binary problem, because the expected result of the procedure is the determination of the presence or not of material in every point of the design space. Unfortunately, it had been shown that solving this problem using a gradient based method leads to unstable solution [28], and, for this reason, a mathematical relaxation is applied to the binary problem: in other words, instead

of imposing a binary relation between the design variables, and the stiffness tensor, a continuous one is defined instead.

For this purpose, the most common strategy is to express the stiffness tensor in function of the density of a structured material. The model of such material must allow the density to cover all the values between 0 and 1, and this is usually done by the mean of microstructural approach: the material is modelled as it was characterized by the repetition of microscopic unit cells, and its behaviour at a macroscopic level depends by the geometry of the cells. Furthermore, the macroscopic properties can be computed by mathematical homogenization methods, and, for this reason, this is referred as homogenization method for topology optimization. Even if many works can be found in literature regarding the theoretical aspects of the homogenization methods, the main point for the present discussion is that these techniques are able to address a relation between a characteristic parameter of the material, or a set of parameters, and the elasticity tensor. The most widely used material interpolation scheme is the so called Solid Isotropic Material with Penalization (SIMP), which impose the following relation between the material density and the elasticity tensor:

$$E_{ijhk} = \left(\rho\left(\underline{x}\right)\right)^{p} \cdot E^{0}_{ijhk} \tag{3.32}$$

It is important to highlight that, again, this is a general statement, which does not involve the finite element method formulation.

A last important remark regards the possibility of avoiding the use of a relaxed formulation: even if the binary problem is not suitable to be solved using gradient based methods, it doesn't mean that it is not possible to solve it at all; later on, an alternative solution methodology will be introduced which allows to solve the actual binary approach, and moreover, its equivalence to gradient based strategies will be addressed.

3.4 Topology Optimization based on the FEA framework and SIMP material interpolation

In the previous sections, the minimum compliance problem has been formulated as follows:

$$
\begin{aligned}
min_{\underline{v}\in U_{adm},E} \quad & \int_{\Omega} \underline{v}^T \cdot \underline{b}\, d\Omega + \int_{\Gamma} \underline{v}^T \cdot \underline{t}\, d\Gamma \\
s.t. \quad & \int_{\Omega} \varepsilon_{ij}(\underline{v})\, E_{ijhk}\, \varepsilon_{hk}(\underline{u}) d\Omega = \int_{\Omega} \underline{v}^T \cdot \underline{b}\, d\Omega + \int_{\Gamma} \underline{v}^T \cdot \underline{t}\, d\Gamma \quad \forall\, \underline{v} \in U_{adm} \\
& \underline{\underline{E}} \in E_{adm}
\end{aligned}
\tag{3.33}
$$

Furthermore, it has been stated that, if a finite element method framework is adopted, and the same mesh is used for both displacements vector and stiffness tensor fields, it may be written:

$$
\begin{aligned}
min_{\underline{u}\in U_{adm},E} \quad & \underline{f}^T \cdot \underline{u} \\
s.t. \quad & \underline{\underline{K}}(E) \cdot \underline{u} = \underline{f} \\
& \underline{\underline{E}} \in E_{adm}
\end{aligned}
\tag{3.34}
$$

where symbols have the usual meanings.

Furthermore, if the SIMP material interpolation law is adopted, the young modulus for the isotropic material can be expressed in function of the element density ρ_e, stored in the vector $\underline{\rho}$. Consequently, the stress/strain matrix, can be written as:

$$
\begin{aligned}
& E = E(\rho_e) = {\rho_e}^k \cdot E_0 \quad \Longrightarrow \quad \underline{\underline{D}} = \frac{E(\rho_e)}{1-v^2} \cdot \begin{vmatrix} 1 & v & 0 \\ v & 1 & 0 \\ 0 & 0 & \frac{1-v}{2} \end{vmatrix} = \\
& = {\rho_e}^k \cdot \frac{E_0}{1-v^2} \cdot \begin{vmatrix} 1 & v & 0 \\ v & 1 & 0 \\ 0 & 0 & \frac{1-v}{2} \end{vmatrix} = {\rho_e}^k \cdot \underline{\underline{D}}_0
\end{aligned}
\tag{3.35}
$$

and the optimization problem may be rewrite as:

$$
\begin{aligned}
min_{\underline{\rho},\underline{u}\in U_{adm},\underline{\rho}} \quad & J = \underline{f}^T \cdot \underline{u}\left(\underline{\rho}\right) \\
s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) - \underline{f} = 0 \\
& h = \textstyle\sum_{e=1}^{n} \rho_e - \overline{V} \leq 0 \\
& \rho_{min} \leq \rho_e \leq 1
\end{aligned}
\tag{3.36}
$$

where, as further constraint, it had been imposed that the total mass of the structure is minor of equal to a certain volume fraction $\overline{V}$.

3.5 Implementation of topology Optimization based on SIMP

3.5.1 Mathematical programming methods

Once the discretization schema (i.e., Finite Elements), the optimization problem (i.e., compliance minimization), and the material characterization (i.e., SIMP) are set up, it is necessary to carry out a numerical procedure in order to solve the optimization problem, which may be written in the synthetic form

$$\begin{aligned} min \quad & J\left(\underline{\rho}\right) = \left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) \\ s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) = \underline{f} \\ & V\left(\underline{\rho}\right) - V_0 = 0 \\ & \rho_{min} \leq \rho_i \leq 1 \quad i = 1, \dots, n \end{aligned} \tag{3.37}$$

and, consequently, determine the topology of the ideal structure. There is a great number of available mathematical programming algorithms which have been implemented in topology optimization, and all of them are gradient based procedures [29].

Let us consider a generic constrained multivariable optimization problem as follows:

$$\begin{aligned} min \quad & J(\underline{x}) \qquad \underline{x} = (x_1, x_2, \dots, x_n)^T \\ s.t. \quad & g_1(\underline{x}) = 0 \\ & g_2(\underline{x}) = 0 \\ & \vdots \\ & g_m(\underline{x}) = 0 \end{aligned} \tag{3.38}$$

The method of the Lagrange multipliers works by defining an augmented Lagrange function

$$\mathcal{L}(\underline{x},\underline{\Lambda}) = \mathcal{L}(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = J(\underline{x}) + \sum_{i=1}^{m} \lambda_i \cdot g_i(\underline{x}) \tag{3.39}$$

and imposing its stationarity as a necessary condition:

$$\begin{aligned} &\left.\frac{\partial \mathcal{L}(\underline{x},\underline{\Lambda})}{\partial \underline{x}}\right|_{\underline{x}=\underline{x}_0} = \frac{\partial J(\underline{x}_0)}{\partial \underline{x}} + \sum_{i=1}^{m} \lambda_i \cdot \frac{\partial g_i(\underline{x}_0)}{\partial \underline{x}} = 0 \\ &\left.\frac{\partial \mathcal{L}(\underline{x},\underline{\Lambda})}{\partial \underline{\Lambda}}\right|_{\underline{x}=\underline{x}_0} = g_i(\underline{x}_0) = 0 \end{aligned} \tag{3.40}$$

where $\underline{x}_0$ is a critical point, such as a maximum, minimum or a saddle.

In order to deal with topology optimization, this formulation must be extended to the case of inequality constraints as well. Let us consider the new formulation:

$$\begin{aligned} &min \quad J(\underline{x}) \qquad \underline{x} = (x_1, x_2, \dots, x_n)^T \\ &s.t. \quad g_1(\underline{x}) \le 0 \\ &\qquad g_2(\underline{x}) \le 0 \\ &\qquad \vdots \\ &\qquad g_m(\underline{x}) \le 0 \end{aligned} \tag{3.41}$$

In this case, it is possible to eliminate the inequality relations by the mean of the slack variables y_i:

$$g_i(\underline{x}) + y_i^2 = 0 \tag{3.42}$$

The Lagrange function depends by the set of variables $\underline{x} = (x_1, x_2, \dots, x_n)^T$, $\underline{y} = (y_1, y_2, \dots, y_m)^T$, and $\underline{\Lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$, and for this reason, it can be written as

$$\begin{aligned} \mathcal{L}(\underline{x},\underline{\Lambda}) &= \mathcal{L}(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m, y_1, y_2, \dots, y_m) \\ &= J(\underline{x}) + \sum_{i=1}^{m} \lambda_i \cdot \left(g_i(\underline{x}) + y_i^2\right) \end{aligned} \tag{3.43}$$

stationarity condition can be written as follows

$$\left.\frac{\partial \mathcal{L}\left(\underline{x},\underline{y},\underline{\Lambda}\right)}{\partial \underline{x}}\right|_{\underline{x}=\underline{x}_0} = \frac{\partial J(\underline{x}_0)}{\partial \underline{x}} + \sum_{i=1}^{m} \lambda_i \cdot \frac{\partial g_i(\underline{x}_0)}{\partial \underline{x}} = 0$$

$$\left.\frac{\partial \mathcal{L}\left(\underline{x},\underline{y},\underline{\Lambda}\right)}{\partial \underline{\Lambda}}\right|_{\underline{x}=\underline{x}_0} = g_i(\underline{x}_0) + y_i^2 = 0 \tag{3.44}$$

$$\left.\frac{\partial \mathcal{L}\left(\underline{x},\underline{y},\underline{\Lambda}\right)}{\partial \underline{y}}\right|_{\underline{x}=\underline{x}_0} = 2 \cdot \lambda_i \cdot y_i = 0$$

It can be noticed that, when a slack variable y_i is null, consequently the i-th Lagrange multiplier λ_i is not null, and, even if $g_i(\underline{x}_0) = 0$, its contribution to the derivative of the Lagrange equation may not be null:

$$\lambda_i \cdot \frac{\partial g_i(\underline{x}_0)}{\partial \underline{x}} \neq 0 \tag{3.45}$$

When the condition $(y_i = 0, \lambda_i \neq 0)$ is satisfied, it is said that the constraint $g_i(\underline{x}) \leq 0 \rightarrow g_i(\underline{x}) = 0$ is active, because its derivative must be actually taken in account: this condition may be intuitively explained by the fact that if $g_i(\underline{x}) < 0$ is strictly negative, its infinitesimal variation doesn't violate the condition despite the direction, and the constraint simply cannot be taken in account; on the contrary, if $g_i(\underline{x}) = 0$, the variation of the objective function is allowed only if it doesn't violate the constrain, which is called active in this case.

Since the determination of the set of the slack variables is not important, they can be simply eliminated by considering the following formulation, called Kurush-Kuhn-Tucker (KKT) conditions:

$$\begin{aligned} &\left.\frac{\partial \mathcal{L}\left(\underline{x},\underline{y},\underline{\Lambda}\right)}{\partial \underline{x}}\right|_{\underline{x}=\underline{x}_0} = \frac{\partial J(\underline{x}_0)}{\partial \underline{x}} + \sum_{i=1}^{m} \lambda_i \cdot \frac{\partial g_i(\underline{x}_0)}{\partial \underline{x}} = 0 \\ &\lambda_i \cdot g_i(\underline{x}_0) + y_i^2 = 0 \\ &\lambda_i \geq 0 \end{aligned} \tag{3.46}$$

The stationarity of the Lagrange equation is the starting point of different optimization algorithms, which have been created in order to deal with different typologies of objective functions and constraints.

Linear programming (LP)

Linear programming is used when objective functions and constraints are linear functions of the design variables:

$$\begin{aligned} min \quad & J(\underline{x}) = c_1 \cdot x_1 + c_2 \cdot x_2 + \cdots + c_n \cdot x_n \\ s.t. \quad & a_{1,1} \cdot x_1 + a_{1,2} \cdot x_2 + \cdots + a_{1,n} \cdot x_n = b_1 \\ & \vdots \\ & a_{m,1} \cdot x_1 + a_{m,2} \cdot x_2 + \cdots + a_{m,n} \cdot x_n = b_m \\ & a_{m+1,1} \cdot x_1 + a_{m+1,2} \cdot x_2 + \cdots + a_{m+1,n} \cdot x_n \leq b_{m+1} \\ & \vdots \\ & a_{m+q,1} \cdot x_1 + a_{m+q,2} \cdot x_2 + \cdots + a_{m+q,n} \cdot x_n \leq b_{m+q} \\ & x_1, x_2, \dots, x_n \geq 0 \end{aligned} \tag{3.47}$$

A possible solution strategy is the adoption of the simplex method, which will be explained in details in the next section, in relation to the evolutionary approach to topology optimization.

Linear programming problems are not so common, especially in studying physical systems, whose models are in general highly not linear. For this reason, Non-Linear Programming procedures has been developed in order to deal with problems such minimization of the compliance of continuum linear elastic structures, or the synthesis of the compliant mechanism. A brief resume of two methods, the sequential quadratic programming and method of moving asymptotes will be reported below, since an extensive description is far from the purposes of the present work.

Non-Linear Programming (NLP) and Sequential Quadratic Programming (SQP)

In general, a quadratic programming problem is formulated as follows:

$$\begin{aligned} min \quad & q(\underline{x}) = \underline{x}^T \cdot \underline{\underline{Q}} \cdot \underline{x} + \underline{c}^T \cdot \underline{x} \\ s.t. \quad & \underline{\underline{A}} \cdot \underline{x} = \underline{b} \\ & \underline{x} \geq 0 \end{aligned} \tag{3.48}$$

In the particular case of the minimum compliance topology optimization of continuum elastic structures, the problem formulation is derived by a quadratic convex approximation of the problem: if $J\left(\underline{\rho}\right)$ is the objective function, then the new formulation

$$\begin{aligned} &min \qquad \underline{x}^T \cdot \underline{\nabla}^2 J\big|_{\underline{\rho}=\underline{\rho}^{(k)}} \cdot \underline{x} + \left(\underline{\nabla} J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}\right)^T \cdot \underline{x} \\ &s.t. \qquad \underline{\underline{A}} \cdot \underline{x} = \underline{0} \end{aligned} \tag{3.49}$$

where $\underline{\nabla} J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}$ and $\underline{\nabla}^2 J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}$ are the gradient and the hessian respectively of the objective function J evaluated in correspondence of the design $\underline{\rho}^{(k)}$, and $\underline{x} = \left(\underline{\rho} - \underline{\rho}^{(k)}\right)$; furthermore, the matrix $\underline{\underline{A}}$ is the matrix representing the constraint linear equation corresponding to the active constraints.

The optimality conditions come from the stationarity of the first derivative of the Lagrange augmented equation:

$$\begin{aligned} &\underline{\nabla}^2 J\big|_{\underline{\rho}=\underline{\rho}^{(k)}} \cdot \underline{x} + \underline{\underline{A}}^T \cdot \underline{\lambda} = -\left(\underline{\nabla} J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}\right)^T \\ &\underline{\underline{A}} \cdot \underline{x} = \underline{0} \end{aligned} \tag{3.50}$$

This approach, referred as active set approach, is conceptually similar to the simplex method, which will be introduced in details later, but whose main idea consists in searching the solution of the optimization problem in the vertices of the hyper polyhedron generated by the union of the hyperplanes representing the linear constraints. The active set approach works in a similar way, despite the fact that the solutions are not necessary in correspondence of the vertices, but, for instance, over an edge between two vertices [30].

The updated design is given by

$$\underline{\rho}^{(k+1)} = \underline{\rho}^{(k)} + \alpha^{(k)} \underline{x}^{(k)} \tag{3.51}$$

where the vector $\underline{x}^{(k)}$ is determined by the solution of the linear system

$$\begin{vmatrix} \left(\underline{\nabla}^2 J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}\right) & \underline{\underline{A}}^T \\ \underline{\underline{A}} & \underline{0} \end{vmatrix} \cdot \begin{vmatrix} \underline{x}^{(k)} \\ \underline{\lambda}^{(k)} \end{vmatrix} = \begin{vmatrix} -\left(\underline{\nabla} J\big|_{\underline{\rho}=\underline{\rho}^{(k)}}\right)^T \\ \underline{0} \end{vmatrix} \tag{3.52}$$

and the value of the scalar $\alpha^{(k)}$ is determined by a line search procedure [28].

Method of Moving Asymptotes

Let us consider the optimization problem P:

$$\begin{aligned} &min && f_0(\underline{x}) \\ &s.t. && f_i(\underline{x}) \leq \bar{\bar{f}}_i && i = 1, \dots, n \\ & && \bar{x}_j \leq x_j \leq \bar{x}_j && j = 1, \dots, n \end{aligned} \tag{3.53}$$

Generally, in structural problems, the objective function f_0 is the total mass of the structure, and the constraint equations $f_i(\underline{x}) \leq \bar{\bar{f}}_i$ represent some limitations on stresses or deformations.

A generalized algorithmic approach for the solution of the problem P consists in generating and solving a set of subproblems $P^{(k)}$ in the context of the following framework:

- set up the iteration step $k = 1$ and the initial guess $\underline{x} = \underline{x}^{(0)}$
- give the iteration step k and the corresponding design $\underline{x}^{(k)}$, compute:
    - the value $f_0(\underline{x}^{(k)}) = f_0^{(k)}$ of the objective function;
    - the value $f_i(\underline{x}^{(k)}) = f_i^{(k)}$ for $i = 1, \dots, n$ of the constraint functions;
    - the gradient $\underline{\nabla} f_i\big|_{\underline{x}=\underline{x}^{(k)}}$ for $i = 0, \dots, n$ for the objective function and constraints;
- generate the subproblem $P^{(k)}$ by the substitution of the generally implicit objective function and constraint functions $f_i(\underline{x})$ for $i = 0, \dots, n$ with a set of explicit functions $f_i^{(k)}(\underline{x})$ generated using the quantities computed in the previous step;
- solve the problem $P^{(k)}$ and set the solution $\underline{x}^{(k+1)}$ as the design for the following iteration step $k + 1$;
- check for convergency, and eventually iterate.

An implementation of this iterative framework has been suggested by Fleury [31], who proposed the determination of $f_i^{(k)}(\underline{x})$ as a linearization of $f_i(\underline{x})$ by the use of a mixed set of variables x_j and $\frac{1}{x_j}$ depending by the sign of the derivative $\frac{\partial f_i}{\partial x_j}$ in correspondence of the design $\underline{x}^{(k)}$. This approach, called CONLIN, states that the following linearization provides a convex approximation of any function $f_i(\underline{x})$:

$$f_i(\underline{x}) = f_i(\underline{x}^{(k)}) + \sum_{+}\left[\left.\frac{\partial f_i}{\partial x_j}\right|_{\underline{x}=\underline{x}^{(k)}} \cdot \left(x_j - x_j^{(k)}\right)\right] + \sum_{-}\left[\left(x_j^{(k)}\right)^2 \cdot \left.\frac{\partial f_i}{\partial x_j}\right|_{\underline{x}=\underline{x}^{(k)}} \cdot \left(\frac{1}{x_j} - \frac{1}{x_j^{(k)}}\right)\right] \tag{3.54}$$

where the symbols $\sum_{+}\cdot$ and $\sum_{-}\cdot$ represent the sum of the terms in correspondence of positive and negative values of the derivative $\frac{\partial f_i}{\partial x_j}$ respectively. This linearization is an approximation of the functions $f_i(\underline{x})$ which owns a simple algebraic structure, and every correspondent subproblem $P^{(k)}$ result convex and separable, and for these reasons they can be easily solved by the application of dual algorithms.

The method of the moving asymptotes (MMA), proposed by Svanberg [32], may be considered a generalization of the CONLIN, because it prescribes a linearization of the functions $f_i(\underline{x})$ using a linear combination of $\frac{1}{x_j - L_j}$ and $\frac{1}{U_j - x_j}$, depending on the sign of the first derivative of f_i. The MMA algorithm follows the iterative framework reported above, which, for its implementation requires the definition of the explicit functions $f_i^{(k)}$ corresponding to the objective function and the constraints. Being $\underline{x}^{(k)}$ the design at the iteration step k, the values L_j and U_j must satisfy the following condition:

$$L_j^{(k)} \leq x_j^{(k)} \leq U_j^{(k)} \tag{3.55}$$

if such condition is satisfied, it is possible to write the convex linearization of $f_i^{(k)}$

$$f_i^{(k)}(\underline{x}) = r_i^{(k)} + \sum_{j=1}^{n}\left[\frac{p_{ij}^{(k)}}{U_j^{(k)} - x_j} + \frac{q_{ij}^{(k)}}{x_j - L_j^{(k)}}\right] \tag{3.56}$$

where

$$p_{ij}^{(k)} = \begin{cases} \left(U_j^{(k)} - x_j^{(k)}\right)^2 \cdot \dfrac{\partial f_i}{\partial x_j} & if \quad \dfrac{\partial f_i}{\partial x_j} > 0 \\ 0 & if \quad \dfrac{\partial f_i}{\partial x_j} \leq 0 \end{cases}$$

$$q_{ij}^{(k)} = \begin{cases} 0 & if \quad \dfrac{\partial f_i}{\partial x_j} \geq 0 \\ -\left(x_j^{(k)} - L_j^{(k)}\right)^2 \cdot \dfrac{\partial f_i}{\partial x_j} & if \quad \dfrac{\partial f_i}{\partial x_j} < 0 \end{cases} \tag{3.57}$$

$$r_i^{(k)} = f_i\left(\underline{x}^{(k)}\right) - \sum_{j=1}^{n}\left[\frac{p_{ij}^{(k)}}{U_j^{(k)} - x_j^{(k)}} + \frac{q_{ij}^{(k)}}{x_j^{(k)} - L_j^{(k)}}\right]$$

Applying the MMA method to the problem of the minimization of the compliance of the continuum structures, it may be noticed that the constraint equations are linear, so the linearization process must be applied only to the objective function. Furthermore, it may be noticed that the derivative of the mean compliance with to respect the density of any element is always negative:

$$\frac{\partial f_0}{\partial x_j} < 0 \quad \forall j = 1, \dots, n \tag{3.58}$$

For these reasons, the formulation of the problem of the compliance, according to the MMA can be expressed as follows [28]:

$$\begin{aligned} min \quad & f_i^{(k)}\left(\underline{x}\right) = -\sum_{j=1}^{n}\left[\frac{\left(x_j^{(k)} - L_j^{(k)}\right)^2}{x_j^{(k)} - L_j^{(k)}} \cdot \frac{\partial f_i}{\partial x_j}\left(\underline{x}^{(k)}\right)\right] \\ s.t. \quad & \sum_{j=1}^{n} x_j^{(k)} \leq \bar{V} \\ & \underline{x} \in \chi^{(k)} \end{aligned} \tag{3.59}$$

where

$$\chi^{(k)} = \left\{\underline{x} \in \chi \;\middle|\; 0{,}9 \cdot L_j^{(k)} + 0{,}1 \cdot x_j^{(k)} \leq x_j \leq 0{,}9 \cdot U_j^{(k)} + 0{,}1 \cdot x_j^{(k)} \;\; i = 1, \dots, n\right\} \tag{3.60}$$

and where $L_j^{(k)}$ and $U_j^{(k)}$ are determined by the heuristic rule proposed by Svanberg [33].

3.5.2 Optimality Criteria

Among others, Optimality Criteria (OC) is one of the most effective procedures employed in topology optimization. This depends by the fact that it is very efficient for the solution of problems with a great number of variables, and few objective functions. In this approach, the solution is obtained by solving the equations deriving by some optimality conditions. The optimality conditions may come from the experience or an intuition, as it appends for the “fully stressed design technique”, or, alternately, an optimality criterion may be the result of stationarity conditions, and may be led back to a gradient based method.

Let us consider the minimum compliance topology optimization problem (3.37)

$$\begin{aligned} min \quad & J\left(\underline{\rho}\right) = \left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) \\ s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) = \underline{f} \\ & V\left(\underline{\rho}\right) - V_0 = 0 \\ & \rho_{min} \leq \rho_i \leq 1 \quad i = 1, \dots, n \end{aligned}$$

then in this case the optimality criterion may be expressed by the following:

$$\begin{cases} if \;\; max\left((1-\zeta)\rho_e{}^{(k)}, \rho_{min}\right) \leq \rho_e{}^{(k)} B_e{}^{(k)} \leq min\left(\left((1+\zeta)\rho_e{}^{(k)}\right), 1\right) & \Rightarrow \;\; \rho_e{}^{(k+1)} = \rho_e{}^{(k)} B_e{}^{(k)} \\ if \;\; \rho_e{}^{(k)} B_e{}^{(k)} < max\left((1-\zeta)\rho_e{}^{(k)}, \rho_{min}\right) & \Rightarrow \;\; \rho_e{}^{(k+1)} = max\left((1-\zeta)\rho_e{}^{(k)}, \rho_{min}\right) \\ \rho_e{}^{(k)} B_e{}^{(k)} > min\left(\left((1+\zeta)\rho_e{}^{(k)}\right), 1\right) & \Rightarrow \;\; \rho_e{}^{(k+1)} = min\left(\left((1+\zeta)\rho_e{}^{(k)}\right), 1\right) \end{cases} \quad (3.61)$$

being $B_e{}^{(k)}$ such that

$$B_e{}^{(k)} = \frac{-\frac{\partial J\left(\underline{\rho}\right)}{\partial \rho_e}}{\lambda \cdot \frac{\partial V\left(\underline{\rho}\right)}{\partial \rho_e}} = -\frac{1}{\lambda} \cdot \left(\underline{u}_e\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \frac{\partial \underline{\underline{K}}_e\left(\underline{\rho}^{(k)}\right)}{\partial \rho_e} \cdot \underline{u}_e\left(\underline{\rho}^{(k)}\right) = -\frac{1}{\lambda} \cdot p\rho_e{}^{p-1}\left(\underline{u}_e\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K}}_0 \cdot \underline{u}_e\left(\underline{\rho}^{(k)}\right) \quad (3.62)$$

where the Lagrange multiplier λ must satisfy the volume constraint, and is actually determined by the means of a bisection algorithm [34]. It can be noticed that, in correspondence of the value $B_e{}^{(k)} = 1$ the updating schema does not produce any variation of the design variables: this means that the state corresponds to a stationary point, and this is coherent with the fact that the KKT conditions lead to the same result, as it will be shown in a while.

Despite the fact that this procedure is probably the most used in topology optimization, in most publications this updating schema appears to have an "heuristic" origin. Nevertheless, it is possible to show that that OC is equivalent to the projected gradient method [35]; to demonstrate this fact, let's start considering the KKT conditions for the optimization problem (3.37), which impose the stationarity of the Lagrange equation:

$$\mathcal{L} = \underline{u}^T \cdot \underline{f} + \lambda\left(\sum \rho_e - \overline{V}\right) + \mu\left(\underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) - \underline{f}\right) - \underline{\alpha}^T \cdot \underline{\rho} + \underline{\beta}^T \cdot \left(\underline{\rho} - \underline{I}\right) \quad (3.63)$$

with to respect the set of variables $\underline{\rho}$ and $\underline{u}$, and the lagrange multipliers λ, μ, $\underline{\alpha}$ and $\underline{\beta}$. Following the procedure proposed by Sigmund [36], it leads to the following:

$$-\left(\underline{u}_e\right)^T \frac{d\underline{\underline{K}}_e\left(\underline{\rho}\right)}{d\rho_e} \underline{u}_e + \lambda - \alpha_e + \beta_e = 0 \quad (3.64)$$

When the bound constraints for the element are not active ($\alpha_e = 0$ and $\beta_e = 0$), the optimality condition in every element with to respect the element density reads as:

$$B_e = \frac{1}{\lambda}\left(\underline{u}_e\right)^T \frac{d\underline{\underline{K}}_e\left(\underline{\rho}\right)}{d\rho_e}\underline{u}_e = 1 \tag{3.65}$$

In order to demonstrate that the previous updating schema corresponds to the optimality condition, let's make an example, considering the simple optimization problem:

$$\begin{aligned} min_{\underline{x}=(x,y)\in\mathbb{R}^2} \quad & J(x,y) = \rho_1^2 + \rho_2^2 \\ s.t. \quad & g(x,y) = \rho_1 + \rho_2 - 1 = 0 \end{aligned} \qquad initial\ guess: \ \underline{\rho} = (-1,2) \tag{3.66}$$

As it is well known, when searching a minimum for a convex function without any constraint, the most efficient updating direction is the one which follows the gradient of the objective function. When the optimization problem is constrained, such constraint must be taken in account. In projected gradient methods, this is done by projecting the gradient of the objective function along the constraint direction, as it is depicted in figure 3.2.

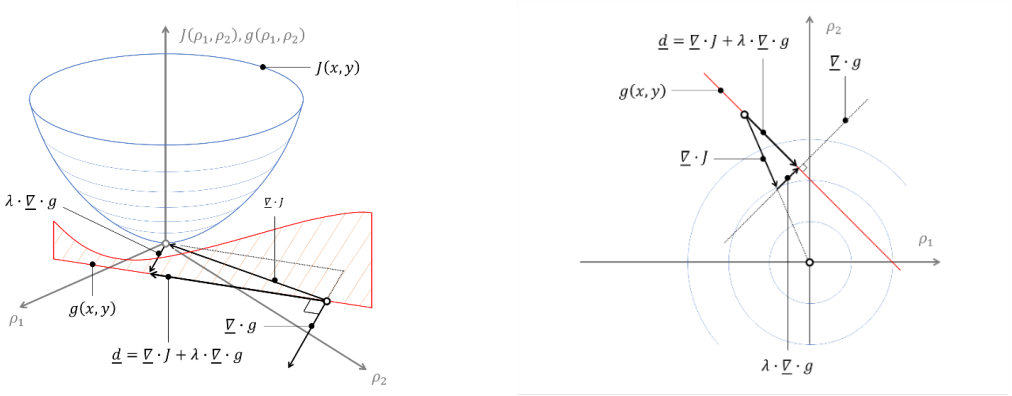
In gradient descent methods, the iteration steps are made along the direction of the gradient of the objective function $\underline{\nabla} \cdot g$, which is the direction of the iteration that leads to the faster convergence of the unconstrained problem. The projected method takes in account the existence of the constraints, and the vector $\underline{d}$ that represents the iteration step, as it is shown in figure, may be expressed as:

$$\underline{d} = -\underline{\nabla} \cdot J + \lambda \cdot \underline{\nabla} \cdot g \quad \Leftrightarrow \quad \lambda = \frac{\underline{\nabla} \cdot f \cdot \underline{\nabla} \cdot g}{\underline{\nabla} \cdot g \ \cdot \underline{\nabla} \cdot g} \tag{3.67}$$

If the iteration step is:

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \gamma \cdot \underline{d} \tag{3.68}$$

it doesn't violate the constraint despite the value of γ. A further important remark is that it is possible to demonstrate that the direction $\underline{d}$ is the one which allows the decreasing of the objective function in the fastest way.



(a) *(b)*

Figure 3.2: example of constrained optimization using the projected gradient method

Returning to the initial problem, it is possible to write the derivative of the objective function with respect to the densities, taking in account the equilibrium equation:

$$\underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) - \underline{f} = 0 \quad \Rightarrow \quad \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) + \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \frac{\partial \underline{u}\left(\underline{\rho}\right)}{\partial \rho_e} = 0 \quad \Rightarrow$$
$$\Rightarrow \frac{\partial \underline{u}\left(\underline{\rho}\right)}{\partial \rho_e} = -\left(\underline{\underline{K}}\left(\underline{\rho}\right)\right)^{-1} \cdot \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) \tag{3.69}$$

Furthermore, the derivative of the objective function is:

$$\frac{\partial}{\partial \rho_e}\left(\left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right)\right) = 2 \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \frac{\partial \underline{u}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) + \left(\underline{u}\left(\underline{\rho}\right)\right)^T \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) =$$
$$= 2 \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \left(-\left(\underline{\underline{K}}\left(\underline{\rho}\right)\right)^{-1} \cdot \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right)\right) \cdot \underline{u}\left(\underline{\rho}\right) + \left(\underline{u}\left(\underline{\rho}\right)\right)^T \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) =$$
$$= -2 \cdot \left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) + \left(\underline{u}\left(\underline{\rho}\right)\right)^T \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) = \tag{3.70}$$
$$= -\left(\underline{u}\left(\underline{\rho}\right)\right)^T \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right) \quad \Rightarrow \frac{\partial J}{\partial \rho_e} = -\left(\underline{u}\left(\underline{\rho}\right)\right)^T \frac{\partial \underline{\underline{K}}\left(\underline{\rho}\right)}{\partial \rho_e} \cdot \underline{u}\left(\underline{\rho}\right)$$

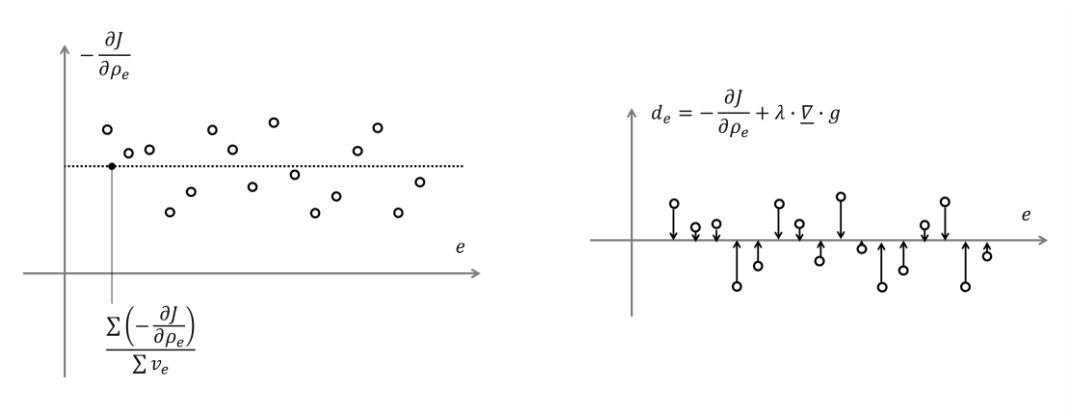
This expression shows that the minimum compliance problem is “self-adjoint”: make reference to Appendix A for the general formulation of the so called adjoint method. Recalling the expression of $g\left(\underline{\rho}\right)$

$$g\left(\underline{\rho}\right) = \sum_{e=1}^{n} \rho_e - \overline{V} \quad \Rightarrow \quad \lambda = \frac{\underline{\nabla} \cdot J \cdot \underline{\nabla} \cdot g}{\underline{\nabla} \cdot g \cdot \underline{\nabla} \cdot g} = \frac{\sum \frac{\partial J}{\partial \rho_e} \cdot \sum v_e}{\sum v_e \cdot \sum v_e} = \frac{\sum \frac{\partial J}{\partial \rho_e} \cdot \sum v_e}{\sum v_e \cdot \sum v_e} = \frac{\sum \frac{\partial J}{\partial \rho_e}}{\sum v_e} \tag{3.71}$$

it may be noticed that, if the volumes of all the elements are identically equivalent to 1, the value of the Lagrange multiplier is equal to the average of the derivatives (sensitivities) of the objective function with to respect the densities of all the elements. Furthermore, if $v_e = 1$ for every element, it means that:

$$\begin{aligned} v_e = 1 \quad e = 1, \dots, n \quad \Longrightarrow \quad & \frac{\partial \underline{\nabla} \cdot g}{\partial \rho_e} = 1 \quad \Longrightarrow \quad \underline{d} = -\underline{\nabla} \cdot J + \lambda \cdot \underline{\nabla} \cdot g \\ & = -\underline{\nabla} \cdot J + \lambda \cdot \underline{\mathrm{I}} \quad \Longrightarrow \\ \Longrightarrow \quad d_e = -\frac{\partial J}{\partial \rho_e} + \lambda & \end{aligned} \tag{3.72}$$

This means that the iteration step for every element represent an increase of the objective function if $\frac{\partial J}{\partial \rho_e}$ is lower than the average λ and vice versa: the interpretation of the behaviour of the iteration step vector $\underline{d}$ is depicted in figure 3-3.



(a) *(b)*

Figure 3.3: behaviour of the components of the vector $\underline{d}$ in the minimum compliance problem

Such behaviour is substantially equivalent to the OC updating schema, noticing that, if, for a particular design variable ρ_e, the value of the sensitivity of the objective function $\frac{\partial J}{\partial \rho_e}$ is minor than λ, where:

$$\lambda = p\underline{\rho}^{p-1} \left(\underline{u}_e \left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K}}_0 \cdot \underline{u}_e \left(\underline{\rho}^{(k)}\right) \tag{3.73}$$

then the value of the variable ρ_e must be increased, otherwise, it must be decreased. In fact, such statement is equivalent to the said updating schema (3.48)

$$\begin{cases} if \;\; max\left((1-\zeta){\rho_e}^{(k)}, \rho_{min}\right) \le {\rho_e}^{(k)} {B_e}^{(k)} \le min\left(\left((1+\zeta){\rho_e}^{(k)}\right), 1\right) & \Rightarrow \;\; {\rho_e}^{(k+1)} = {\rho_e}^{(k)} {B_e}^{(k)} \\ if \;\; {\rho_e}^{(k)} {B_e}^{(k)} < max\left((1-\zeta){\rho_e}^{(k)}, \rho_{min}\right) & \Rightarrow \;\; {\rho_e}^{(k+1)} = max\left((1-\zeta){\rho_e}^{(k)}, \rho_{min}\right) \\ {\rho_e}^{(k)} {B_e}^{(k)} > min\left(\left((1+\zeta){\rho_e}^{(k)}\right), 1\right) & \Rightarrow \;\; {\rho_e}^{(k+1)} = min\left(\left((1+\zeta){\rho_e}^{(k)}\right), 1\right) \end{cases}$$

3.6 Evolutionary approach for topology optimization

It may be recalled that the design technique based on the SIMP material interpolation allows a relaxation of the binary optimization problem: it means that, instead of determining if material exists or not in correspondence of a certain point, or element, continuous computation of the densities is provided. Furthermore, SIMP doesn't require any remeshing process. Anyway, there is another set of TO techniques which avoids remeshing, but keeps the binary nature of the problem: they are the evolutive techniques (ESO, BESO, AESO, Soft Kill BESO, etc…).

Historically, the first one has been ESO, which is based on the observation of some biological tissues [37]. In fact, living being structures, such as bones and shells, are generated by an evolutionary process which takes a long time and depends on the surrounding environment conditions. Under an evolutionary point of view, structural topology optimization tries to mimic such behaviour. The methodology is based on the same finite element framework introduced in the previous sections: the design domain, the constraints and the external forces are defined, and a feasible discretization of the continuum is provided by the use of 4 nodes isoperimetric mesh.

At a very beginning of the procedure, the design domain is supposed to be full of material, and the first FE analysis is carried out. Once the displacements, strain and stress fields are determined, it is possible to apply the evolutionary process, which consists in eliminating from the mesh some elements on the base of a certain rejection criterion (RC). For example, elements are deleted when the Von Mises stress is less than a rejection ratio (RR) times the maximum allowed over the structure. Once the structure is defined, it is possible to carry out a new FE analysis, and start with another iteration step.

As it appended for the optimality criteria applied to compliance minimization in the previous section, it is possible to demonstrate that, even if the evolutionary approach has been inspired by a heuristic observation, it can be led back to a gradient based method

[38]. As a relevant remark, the reasoning starts from the recall of the Michell truss theory, which may be considered the very first example in the structural optimization field.

3.6.1 Theoretical foundation of the evolutionary approach

Michell, in his famous article [39], has been probably the first in formally formulating a topology optimization problem. More in detail, his methodology allows the determination of the frame structure able to fulfil resistance requirements, and having the minor possible volume.

The reasoning of Michell starts from another relevant result from Maxwell [40], who stated that in every frame subject to a certain system of forces, the sum of the lengths of the beams subject to compression multiplied by the internal forces they are subject to, minus the sum of the lengths of the beams subject to traction multiplied by the internal forces they are subject to, is a constant. In the symbolic formulation recalled by Michell:

$$\sum f_p \cdot l_p - \sum f_q \cdot l_q = C \tag{3.74}$$

Starting from this assumption, Michell derived two relevant results: firstly, he stated that, for a certain limit of allowed stress, the truss with the lower amount of material is the one having all the same load in all its parts (in absolute value). Furthermore, he disclosed that the minimum volume trusses are composed by a net of elements having a particular topology, which may be described by the so-called Hencky nets [41]. As it will be highlighted later on, one relevant property of this structures is that the direction of their elements is coherent with the orientation of the eigenvectors of the stress tensor (principal directions).

Definition of the objective function

Among other considerations, there is a property of the Michell truss which agrees with ESO procedure described above: in fact, Michell truss is the structure showing the minimum ratio between volume and stiffness, or, equivalently, the minimum volume times compliance product $W \cdot V$. With reference to figures 3.4, it is possible to show that the product $W \cdot V$ is a feasible objective function, in order to translate the optimal criterion derived by the Michell statements in a mathematical formulation. As a remark, it is

highlighted here that the example is visualized using four ground structures, but the reasoning is perfectly equivalent in the case of continuum structures.

Let us consider the two ground structures depicted in figure 3.4 (a) and in figure 3.4 (b): they are both subject to a unit, or dummy, vertical load, and their total volume is V_1; if the right-side structure is a generic truss, and the left side one is the corresponding Michell truss; then, according to the optimality condition for it is possible to write

$$W_{M1} \leq W_{T1} \quad \Rightarrow \quad W_{M1} \cdot V_1 \leq W_{T1} \cdot V_1 \tag{3.75}$$

where W_{M1} and 1 are the values of the corresponding compliances for the two trusses.

On the other hand, let us consider other two ground structures, depicted in figure 3.4(c) and figure 3.4(d): a generic truss on the right side, and a Michell truss on the left side one; both are subject to the previous load conditions, and they have the same volume V_2. Under these assumptions, it is possible to write:

$$W_{M2} \leq W_{T2} \quad \Rightarrow \quad W_{M2} \cdot V_2 \leq W_{T2} \cdot V_2 \tag{3.76}$$

Let us now consider the two Michell trusses, the one of volume V_1 in figure 3.4 (a), and the one of volume V_2 in figure 3.4 (c). It is possible notice that they have the same topology (determined by the geometry of the Hencky net corresponding to the load conditions); furthermore, the two following statements are true for the structures:

- if all the cross sections of the elements of the Michell truss of volume V_1 are multiplied by the quantity V_1/V_2, the volume of the new structure is V_2 (this sentence is true by elementary geometric considerations);
- in the hypothesis of small displacements, if the material is linear elastic, and if all the cross sections of the volume V_1 truss are multiplied by V_2/V_1, then the resulting compliance is multiplied by the quantity V_2/V_1 as well (this is true by the linearity of the linear elastic problem with small displacements).

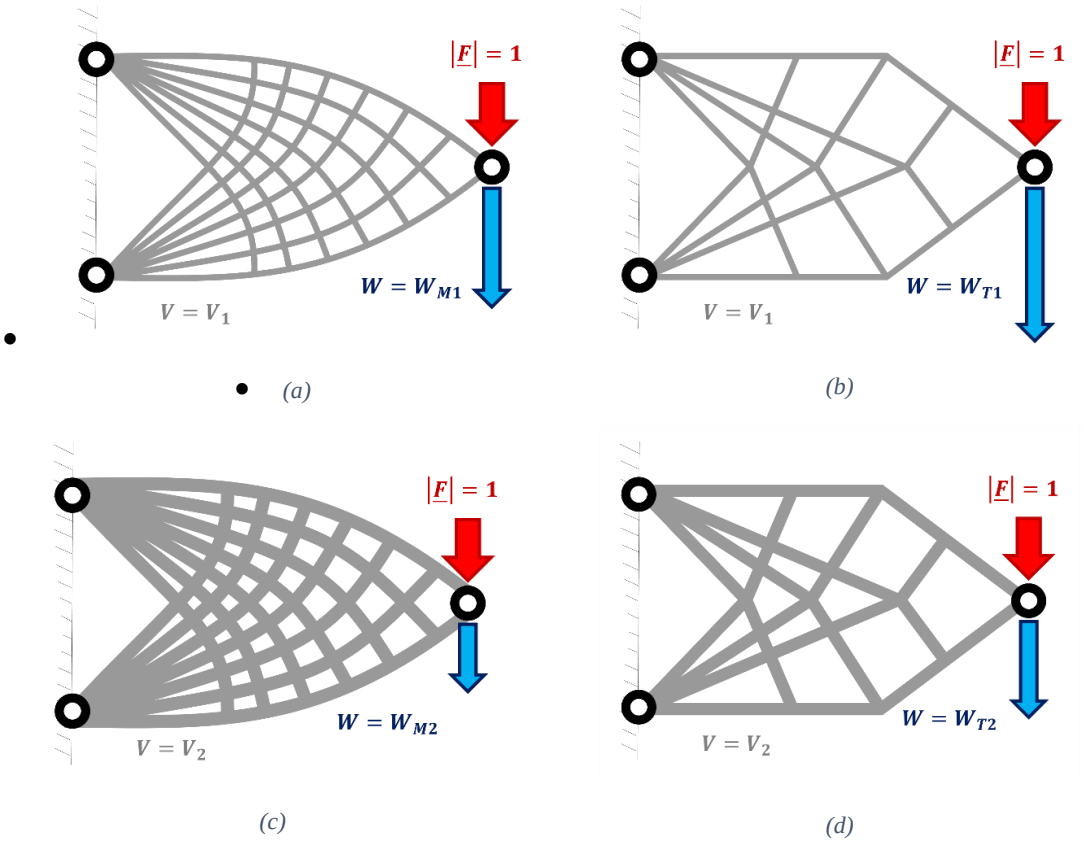


Figure 3.4: different ground structures characterized by different values of the product W·V

Consequently, it is possible to determine the following relation between the two systems:

$$\left.\begin{aligned} V_2 &= \frac{V_2}{V_1}\cdot V_1 \\ W_{M2} &= \frac{V_1}{V_2}\cdot W_{M1} \end{aligned}\right\} \quad \Rightarrow \quad W_{M2}\cdot V_2 = \frac{V_1}{V_2}\cdot W_{M1}\cdot\frac{V_2}{V_1}\cdot V_1 = W_{M1}\cdot V_1 \tag{3.77}$$

Then, on the base of the inequality, it is possible to write:

$$W_{M2}\cdot V_2 \le W_{T2}\cdot V_2 \quad \Rightarrow \quad W_{M1}\cdot V_1 \le W_{T2}\cdot V_2 \tag{3.78}$$

Verbally speaking, the last inequality means that, independently by the volume, or by the deflection of the application point of the dummy load, a Michell truss has always an inferior or equal ratio compliance/volume, compared to any other possible truss.

The formulation of NLP problem and its linearization

Once the objective function has been defined, it is possible to define the constrained optimization problem, that in the case of a continuum structure, reads as follows:

$$\begin{aligned} &min \quad \tilde{J}\left(\underline{\rho}\right) = W\left(\underline{\rho}\right) \cdot V\left(\underline{\rho}\right) \\ &s.t. \quad 0 \leq \rho_i \leq \rho_{max} \quad i = 1, \dots, n \end{aligned} \tag{3.79}$$

being $\underline{\rho}$ the vector containing the densities of the finite elements.

The previous minimization problem has a logarithmic equivalent form [42]:

$$\begin{aligned} &min \quad \tilde{J}\left(\underline{\rho}\right) = \ln\left(W\left(\underline{\rho}\right) \cdot V\left(\underline{\rho}\right)\right) \quad \Rightarrow \\ &s.t. \quad 0 \leq \rho_i \leq \rho_{max} \quad i = 1, \dots, n \\ \Rightarrow \quad &min \quad J\left(\underline{\rho}\right) = \ln\left(W\left(\underline{\rho}\right)\right) + \ln\left(V\left(\underline{\rho}\right)\right) \\ &s.t. \quad 0 \leq \rho_i \leq \rho_{max} \quad i = 1, \dots, n \end{aligned} \tag{3.80}$$

This is a non-linear programming problem, and one feasible solution techniques are the sequential linear programming algorithms. Briefly, these algorithms provide a sequential linearization of the nonlinear problem, in order to be able to apply linear programming strategies.

The linearization of the objective function is obtained as follows:

$$J\left(\underline{\rho}\right) \cong J\left(\underline{\rho}^{(k)}\right) + \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \left(\underline{\rho} - \underline{\rho}^{(k)}\right) \tag{3.81}$$

where $\underline{\rho}^{(k)}$ is the point at which the linearization is caried out (at a certain iteration step k), and the gradient elements of the function J calculated in the point $\underline{\rho}^{(k)}$ are:

$$\frac{dJ\left(\underline{\rho}^{(k)}\right)}{d\rho_i} = \frac{1}{V\left(\underline{\rho}^{(k)}\right)} \cdot A_i \cdot \left(1 - \frac{w_i\left(\underline{\rho}^{(k)}\right)}{w_0\left(\underline{\rho}^{(k)}\right)}\right) \tag{3.82}$$

where $w_i\left(\underline{\rho}^{(k)}\right)$ is the average strain energy density in the i-th element, and $w_0\left(\underline{\rho}^{(k)}\right)$ is the average strain energy in the structure. This expression states that al the sensitivities (derivatives) of the objective function with respect to the element densities are all equal to zero (stationary point), if the strain energy densities of the elements are all equal to the average of the global strain energy density.

An interesting remark is that actually this expression has the same interpretation of the updating formula seen for the OC for the minimization of the strain energy with SIMP material characterization. Recalling what has been stated in the previous sections it is possible to write:

$$\left.\begin{aligned} \lambda &= \frac{\sum \frac{\partial J}{\partial \rho_e}}{\sum v_e} \\ d_e &= -\frac{\partial J}{\partial \rho_e} + \lambda \end{aligned}\right\} \quad \Rightarrow \quad d_e = -\frac{\partial J}{\partial \rho_e} + \frac{\sum \frac{\partial J}{\partial \rho_e}}{\sum v_e} = 0 \quad \Leftrightarrow \quad \frac{\partial J}{\partial \rho_e} = \frac{\sum \frac{\partial J}{\partial \rho_e}}{\sum v_e} \tag{3.83}$$

expression which, again, states the equality between the elements strain energy, and the average value. This is an evidence of the fact that the behaviour of the solution of the minimum compliance problem is independent by the type of formulation involved.

The solution of the LP problem

The equivalent optimization problem is formulated as follows

$$\begin{aligned} min \quad & J\left(\underline{\rho}^{(k)}\right) + \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \left(\underline{\rho} - \underline{\rho}^{(k)}\right) \\ s.t. \quad & \frac{dJ\left(\underline{\rho}^{(k)}\right)}{d\rho_i} = \frac{1}{V\left(\underline{\rho}^{(k)}\right)} \cdot A_i \cdot \left(1 - \frac{w_i\left(\underline{\rho}^{(k)}\right)}{w_0\left(\underline{\rho}^{(k)}\right)}\right) \\ & 0 \le \rho_i \le \rho_{max} \quad i = 1, \dots, n \end{aligned} \tag{3.84}$$

which, according to what has been introduced in the previous section, is a linear programming problem indeed. In fact, a standard LP formulation is defined by:

- a set of variables $\underline{x} = (x_1, \dots, x_n)^T$;
- an objective function $J(\underline{x}) = \underline{c} \cdot \underline{x}^T$;
- a set of constraints $\underline{\underline{A}} \cdot \underline{x} \le \underline{b}$

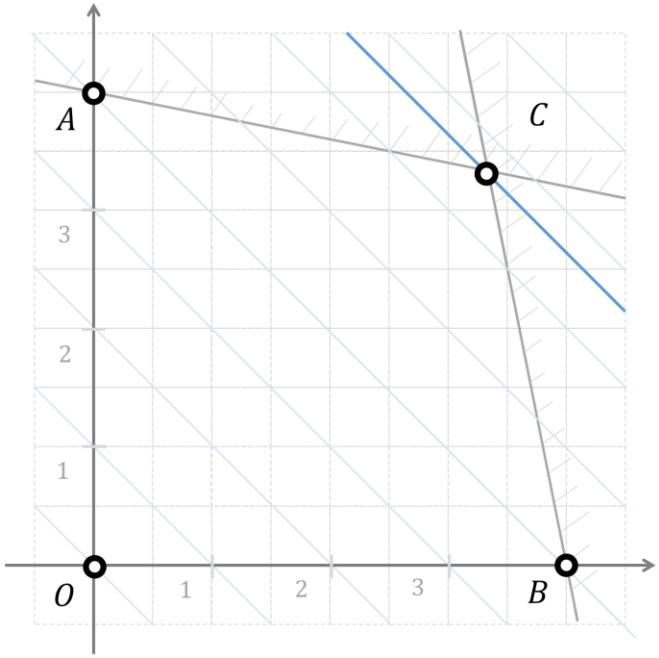
An important property of the LP problems is that it is always possible to define a dual problem:

$$\begin{aligned} min \quad & g(\underline{x}) = \underline{b} \cdot \underline{y}^T \\ s.t. \quad & \left(\underline{\underline{A}}\right)^T \cdot \underline{y} \le \underline{c} \end{aligned} \tag{3.85}$$

equivalent to the problem:

$$\begin{aligned} min \quad & f(\underline{x}) = \underline{c} \cdot \underline{x}^T \\ s.t. \quad & \underline{\underline{A}} \cdot \underline{x} \le \underline{b} \end{aligned} \tag{3.86}$$

As it is stated, the optimization is a so-called bounded LP problem, which are characterized by some features: as a first remark, it is a convex optimization problem, and furthermore, it is known that the global minimum (maximum) is placed in correspondence of one of the vertices of the hyper polyhedron defined by the inequality constraints.



Problem statement:

$f(\underline{x}) = f(x_1, x_2) = x_1 + x_2$
$x_1 + 4 \cdot x_2 \leq 16$
$4 \cdot x_1 + x_2 \leq 16$
$x_1 \geq 0$
$x_2 \geq 0$

Optimal solution (max):

$f(\underline{x}) = \frac{8}{5}$
$x_1 = \frac{16}{5}$
$x_2 = \frac{16}{5}$

Optimal solution (min):

$f(\underline{x}) = 0$
$x_1 = 0$
$x_2 = 0$

Figure 3.5: graphic representation of a linear programming optimization problem

To visually illustrate the last sentence, figure 3.5 reports the graphical solution of a simple two dimensions linear programming problem: actually, the minimum for the objective function is placed in correspondence of the origin of the reference system, the point O, and the maximum in the point C, which are two vertices of the domain defined by the constraint inequalities.

This idea is the base of the simplex resolution method: this procedure starts from a feasible state of the system, represented by a vertex of the domain, and update the state variables moving to an adjacent vertex, according to the values of the coefficients of the objective function.

In order to apply the simplex method, it is necessary to write the problem in a standard form, and this is done by applying these rules:

1) if the problem is a minimization, it must be converted to a maximization problem;
2) all constraints must be grater or equal to zero; if not, the variable x_i must be replaced by $x_i' - x_i''$ with $x_i' \geq 0$ and $x_i'' \geq 0$;

3) if any equality constraint $x_i = b_i$ is present, it must be spitted in two inequality constraints, $x_i \leq b_i$ and $x_i \geq b_i$.
4) Inequalities must be in the form of the kind: $a_{i1} \cdot x_1 + a_{i2} \cdot x_2 + \cdots + a_{in} \cdot x_n = b_i$.

Let us return to the original the topology optimization problem (3.71)

$$\begin{aligned} min \quad & J\left(\underline{\rho}^{(k)}\right) + \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \left(\underline{\rho} - \underline{\rho}^{(k)}\right) \\ s.t. \quad & \frac{dJ\left(\underline{\rho}^{(k)}\right)}{d\rho_i} = \frac{1}{V\left(\underline{\rho}^{(k)}\right)} \cdot A_i \cdot \left(1 - \frac{w_i\left(\underline{\rho}^{(k)}\right)}{w_0\left(\underline{\rho}^{(k)}\right)}\right) \\ & 0 \leq \rho_i \leq \rho_{max} \quad i = 1, \dots, n \end{aligned}$$

then, it is possible to rewrite the objective function:

$$\begin{aligned} & J\left(\underline{\rho}\right) \cong J\left(\underline{\rho}^{(k)}\right) + \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \left(\underline{\rho} - \underline{\rho}^{(k)}\right) = \\ & = \left(J\left(\underline{\rho}^{(k)}\right) - \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\rho}^{(k)}\right) + \left(\underline{\nabla} \cdot J\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\rho} = \\ & = C + \underline{c}^T \cdot \underline{\rho} \end{aligned} \tag{3.87}$$

where the constant C can be ignored in order to carry out the optimization.

The constraints are non-negative, so that, according to the second rule for the canonical form, it is necessary the introduction of n slack variables S_i, so that the constraints are modified as follows:

$$\begin{aligned} & \rho_1 + S_1 = \rho_{max} \qquad \rho_1 \geq 0 \qquad S_1 \geq 0 \\ & \rho_2 + S_2 = \rho_{max} \qquad \rho_2 \geq 0 \qquad S_2 \geq 0 \\ & \qquad \vdots \\ & \rho_n + S_n = \rho_{max} \qquad \rho_n \geq 0 \qquad S_n \geq 0 \end{aligned} \tag{3.88}$$

A feasible initial solution is $\rho_i = \rho_{max}\ \forall i$, (that correspond to the work space full of material), and, consequently, $S_i = 0\ \forall i$. The variables equal to ρ_{max} are called base-variables, and the ones equal to zero are non-basic variables.

The objective function write takin in account the slack variables is:

$$J\left(\underline{\rho}\right) \cong \underline{c}^T \cdot \underline{\rho} = \sum_{i=1}^{n} c_i \cdot (\rho_i - S_i) \tag{3.89}$$

The updating schema must be able to produce the maximum possible increment of this function, and it appends if the generic quantity $(\rho_i - S_i)$, corresponding to the higher coefficient c_i, is transformed in a basic variable. Furthermore, it may be recalled that the generic coefficient c_i is:

$$c_i = \frac{df\left(\underline{\rho}^{(k)}\right)}{d\rho_i} = \frac{1}{V\left(\underline{\rho}^{(k)}\right)} \cdot A_i \cdot \left(1 - \frac{w_i\left(\underline{\rho}^{(k)}\right)}{w_0\left(\underline{\rho}^{(k)}\right)}\right) \tag{3.90}$$

The procedure turns to base variables the i-th term $(\rho_i - S_i)$ having the higher sensitivity value, which means that the variable ρ_i is turned from the value ρ_{max} to the value 0, and it appends in correspondence of the elements having the minor value of strain energy density $w_i\left(\underline{\rho}^{(k)}\right)$. Actually, this is exactly what is prescribed by the optimality criterion in ESO.

There is a last important remark about the equivalence between ESO and the application of the OC to the minimum compliance problem with SIMP material interpolation: despite the fact that the application of the former provides a discrete variation of the design variables, meanwhile the latter has a continuous variation of the element densities, the two approaches are substantially equivalent: the only difference is that, in the ESO approach as it had been just presented, there is not a constraint on the required volume fraction. Nevertheless, in the next section, it will be presented a modified procedure having this feature as well.

3.6.2 Bidirectional evolutionary approach

It had been shown in the previous section that the simplex method allows the recovery of an element eliminated by the previous optimization steps. This behaviour was not included in the very original implementation of the ESO approach, which had been included in the so called BESO (Bidirectional Evolutionary Structural Optimization) algorithms later on. Actually, the latter is the one the one which have been used in this research to test the biomimetic directional paradigm.

Compared to the original version of ESO, the current BESO method [43] has a clear mathematical statement

$$\begin{aligned} min \quad & J\left(\underline{\rho}\right) = \frac{1}{2}\left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) \\ s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) = \underline{f} \\ & V\left(\underline{\rho}\right) - V_0 = 0 \\ & \rho_i = \begin{cases} 1 \\ \rho_{min} \end{cases} \quad i = 1, \dots, n \end{aligned} \tag{3.91}$$

It is interesting confront this formulation with the formulation provided by the use of the SIMP material interpolation:

$$\begin{aligned} min \quad & J\left(\underline{\rho}\right) = \left(\underline{u}\left(\underline{\rho}\right)\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) \\ s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) = \underline{f} \\ & V\left(\underline{\rho}\right) - V_0 = 0 \\ & \rho_{min} \leq \rho_i \leq 1 \quad i = 1, \dots, n \end{aligned} \tag{3.92}$$

which are quite similar despite the factor $\frac{1}{2}$ in the objective function (that actually doesn't affect the result of the optimization process), and the constraint on the state variables.

This similarity comes from the fact, similarly to SIMP, even BESO adopts the same relaxation of the variable expressing the relation (3.32) between the material characterization and the Young's modulus of the material:

$$E_{ijhk} = \left(\rho(\underline{x})\right)^p \cdot E^0_{ijhk}$$

In this way, the sensitivity of the objective function to the variation of the e-th state variable is:

$$\frac{\partial J\left(\underline{\rho}\right)}{\partial x \rho_e} = -\frac{1}{2} \cdot \underline{\rho}^{p-1} \left(\underline{u_e}\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K_0}} \cdot \underline{u_e}\left(\underline{\rho}^{(k)}\right) \tag{3.93}$$

so that the sensitivity number for the so called soft-kill BESO algorithm become:

$$-\frac{1}{p} \cdot \frac{\partial J\left(\underline{\rho}\right)}{\partial x_e} = \begin{cases} \frac{1}{2} \cdot \left(\underline{u_e}\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K_0}} \cdot \underline{u_e}\left(\underline{\rho}^{(k)}\right) & if\ \rho_e = 1 \\ \frac{1}{2} \cdot \underline{\rho}^{p-1} \left(\underline{u_e}\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K_0}} \cdot \underline{u_e}\left(\underline{\rho}^{(k)}\right) & \rho_e = \rho_{min} \end{cases} \tag{3.94}$$

or, imagining to set $p \to \infty$, we obtain the the sensitivity number for the so called hard-kill BESO algorithm

$$-\frac{1}{p}\cdot\frac{\partial J\left(\underline{\rho}\right)}{\partial x_e}=\begin{cases}\frac{1}{2}\cdot\left(\underline{u}_e\left(\underline{\rho}^{(k)}\right)\right)^T\cdot\underline{\underline{K}}_0\cdot\underline{u}_e\left(\underline{\rho}^{(k)}\right) & if\ \rho_e=1\\ 0 & \rho_e=\rho_{min}\end{cases} \tag{3.95}$$

Furthermore, as it will be shown in the next sections, the BESO algorithm adopts a bisection algorithm, in order to update the design variables: similarly to what appends in SIMP, this computation procedure is adopted in order to trying to obtain the same level of strain energy in all the actual elements of the structure.

3.7 Other theoretic features in Continuum Topology Optimization

It is now clear that the synthesis of structures using the topology optimization is a complex task. In fact, it involves the use of tools based on different theoretical backgrounds: for instance, under an engineering point of view, it is necessary to provide an efficient material interpolation schema; speaking about the mathematical features, the setup of the optimization problems has a primary role on the quality of the solution obtained; furthermore, even the computational issues are of great relevance for stability and convergence.

Even if, in the previous sections, the main ingredients of some of the most popular topology optimization algorithms have been introduced, some highlights on more specific, but really important aspects will be provided below.

3.7.1 Convexity of the minimum compliance problem

As it had been largely discussed in this chapter, all the reported solution algorithms for the minimization of the compliance, or, under another point of view, the minimization of the total strain energy, are explicit gradient based methods, or are based on optimality criteria which are equivalent to a gradient based method.

One of the main issues in optimization is to ensure that, once obtained a solution in correspondence of a stationary point, actually this solution is global minimum of the objective function in the studied domain. In unconstrained optimization, the property of convexity of the objective function is a sufficient condition in order to ensure to reach a global minimum. Similarly, in constrained optimization, if both objective functions and

the sub domain defined by the constrains are convex, again the reached solution is a global minimum [44][45].

Convex optimization is an important topic, with applications in mathematic of course, physic, engineering, economy, etc… and provide even a brief resume of the main concepts is far beyond the scope of the present work. Nevertheless, Svanberg [46] disclosed that the problem of minimizing the compliance of a structure is actually a convex optimization problem, and thus, its solution leads to a global minimum.

The convexity issue will be addressed in chapter 5 as well, talking about the design of the compliant mechanisms. We will see that one technique for the synthesis of these particular devices is topology optimization: unfortunately, in this case, the objective function implemented are not convex, therefore leading to a non-global solution.

3.7.2 Confront between strain energy criteria and resistance criteria

In the previous sections it had been demonstrated the substantial equivalence between the gradient based optimization procedures for the minimization of the compliance adopting the SIMP interpolation schema, and the evolutive structural optimization method ESO. Actually, the proof of this equivalence is based on the fact that, as optimality criterion (rejection criterion) for the ESO has been taken in account the evaluation of the strain energy density of the elements. Anyway, historically, the very first rejection criterion was based on the evaluation of the resistance measure, such as the evaluation of the Von Mises equivalent stress: in this case the optimality condition consists in neglecting the elements characterized by a state of tension characterized by a low distortion energy (low equivalent Von Mises Stress).

Nevertheless, it is possible to demonstrate that, in order to obtain the solution structure, the stress criterion is equivalent to the stiffness criterion [47]. To do so, it is necessary to demonstrate that the order in which the elements are neglected in the two cases is the same.

As it had been already stated, the minimum compliance problem is based on the evaluation of the sensitivity of the objective function to the variation of the state variables, that, for every finite element we found it is:

$$\alpha_s = \left(\underline{u}_e\left(\underline{\rho}^{(k)}\right)\right)^T \cdot \underline{\underline{K}}_0 \cdot \underline{u}_e\left(\underline{\rho}^{(k)}\right) \tag{3.96}$$

Recalling that

$$\underline{\underline{K}}_E = t\sum_i\sum_j w_i w_j \left(\underline{\underline{B}}(\xi,\eta)\right)^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}}(\xi,\eta) \tag{3.97}$$

it is possible to write

$$\alpha_s = \left(t\sum_i\sum_j w_i w_j \cdot det\left(\underline{\underline{J}}(\xi,\eta)\right)\right) \cdot \left(\underline{u}_e{}^T \cdot \underline{\underline{B}}^T \cdot \underline{\underline{D}} \cdot \underline{\underline{B}} \cdot \underline{u}_e\right) \tag{3.98}$$

where term in parenthesis

$$\left(t\sum_i\sum_j w_i w_j \cdot det\left(\underline{\underline{J}}(\xi,\eta)\right)\right) = c \tag{3.99}$$

is actually a constant value. Furthermore, recalling the equation (3.18),

$$\underline{\varepsilon} = \begin{vmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \gamma_{xy}\end{vmatrix} = \underline{\underline{B}} \cdot \underline{u}_E \tag{3.100}$$

and the equation (3.21)

$$\begin{aligned}\underline{\sigma} &= \begin{vmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \tau_{xy}\end{vmatrix} = \frac{E}{1-\upsilon^2}\cdot\begin{vmatrix}1 & \upsilon & 0\\ \upsilon & 1 & 0\\ 0 & 0 & \frac{1-\upsilon}{2}\end{vmatrix}\cdot\begin{vmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \gamma_{xy}\end{vmatrix} = \underline{\underline{D}}\cdot\underline{\varepsilon} \quad \Longrightarrow\\ \underline{\varepsilon} &= \left(\underline{\underline{D}}\right)^{-1}\cdot\underline{\sigma} = \left(\frac{E}{1-\upsilon^2}\right)^{-1}\cdot\begin{vmatrix}\frac{1}{(1-\upsilon^2)} & \frac{-\upsilon}{(1-\upsilon^2)} & 0\\ \frac{-\upsilon}{(1-\upsilon^2)} & \frac{1}{(1-\upsilon^2)} & 0\\ 0 & 0 & \frac{2}{(1-\upsilon)}\end{vmatrix}\cdot\underline{\sigma}\end{aligned} \tag{3.101}$$

finally, it is possible to write:

$$\begin{aligned}&\alpha_s = c\cdot\left(\underline{u}_e{}^T\cdot\underline{\underline{B}}^T\cdot\underline{\underline{D}}\cdot\underline{\underline{B}}\cdot\underline{u}_e\right) = c\cdot\left(\underline{\epsilon}^T\cdot\underline{\underline{D}}\cdot\underline{\varepsilon}\right) =\\ &= c\cdot\left(\underline{\sigma}^T\cdot\left(\underline{\underline{D}}^{-1}\right)^T\cdot\underline{\underline{D}}\cdot\underline{\underline{D}}^{-1}\cdot\underline{\sigma}\right) = c\cdot\left(\underline{\sigma}^T\cdot\underline{\underline{D}}^{-1}\cdot\underline{\sigma}\right)\end{aligned} \tag{3.102}$$

On the other hand, the equivalent Von Mises stress is:

$$\sigma_{VM} = \sqrt{{\sigma_{xx}}^2 + {\sigma_{yy}}^2 - \sigma_{xx} \cdot \sigma_{yy} + 3\tau_{xy}} \tag{3.103}$$

For the determination of the order of elimination in the ESO approach, instead of takin in account the quantity σ_{VM}, equivalently, it is possible to take in account its square:

$$\begin{aligned} &{\sigma_{VM}}^2 = {\sigma_{xx}}^2 + {\sigma_{yy}}^2 - \sigma_{xx} \cdot \sigma_{yy} + 3\tau_{xy} = \\ &= \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix}^T \cdot \begin{vmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 3 \end{vmatrix} \cdot \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix} = \underline{\sigma}^T \cdot \underline{\underline{T}} \cdot \underline{\sigma} \end{aligned} \tag{3.104}$$

It is important to recall that this expression of the Von Mises equivalent stress is a pointwise formulation: in order to take in account that this quantity is an average over all the finite element, again, this is calculated by the mean of the Gauss points stress:

$${\sigma_{VM}}^2 = \left(t \sum_i \sum_j w_i w_j \cdot det\left(\underline{\underline{J}}(\xi, \eta) \right) \right) \cdot \left(\underline{\sigma}^T \cdot \underline{\underline{T}} \cdot \underline{\sigma} \right) = c \cdot \left(\underline{\sigma}^T \cdot \underline{\underline{T}} \cdot \underline{\sigma} \right) \tag{3.105}$$

Resuming, we have two quadratic forms of the stress state:

$$\begin{aligned} &\Phi(\underline{\sigma}) = \Phi(\sigma_{xx}, \sigma_{yy}, \tau_{xy}) = \frac{\alpha_s}{c} = \underline{\sigma}^T \cdot \underline{\underline{D}}^{-1} \cdot \underline{\sigma} = \\ &= \frac{1}{E} \cdot \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix}^T \cdot \begin{vmatrix} 1 & -\upsilon & 0 \\ -\upsilon & 1 & 0 \\ 0 & 0 & 2 \cdot (1 + \upsilon) \end{vmatrix} \cdot \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix} \\ &\Psi(\underline{\sigma}) = \Psi(\sigma_{xx}, \sigma_{yy}, \tau_{xy}) = \frac{{\sigma_{VM}}^2}{c} = \underline{\sigma}^T \cdot \underline{\underline{T}} \cdot \underline{\sigma} = \\ &= \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix}^T \cdot \begin{vmatrix} 1 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 3 \end{vmatrix} \cdot \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix} \end{aligned} \tag{3.106}$$

In order to demonstrate the equivalence between the two optimality criteria defined by the elimination of the elements based on the two quadratic forms, it is necessary to

demonstrate that, given two generic states of tension, $\underline{\sigma_A} = (\sigma_{Axx}, \sigma_{Ayy}, \tau_{Axy})$, and $\underline{\sigma_B} = (\sigma_{Bxx}, \sigma_{Byy}, \tau_{Bxy})$, then only the following cases are possible:

$$\begin{array}{lllll} a) & \Phi(\underline{\sigma_A}) > \Phi(\underline{\sigma_B}) & \Rightarrow & \Psi(\underline{\sigma_A}) > \Psi(\underline{\sigma_B}) & \\ b) & \Phi(\underline{\sigma_A}) = \Phi(\underline{\sigma_B}) & \Rightarrow & \Psi(\underline{\sigma_A}) = \Psi(\underline{\sigma_B}) & \forall \underline{\sigma_A}, \underline{\sigma_B} \\ c) & \Phi(\underline{\sigma_A}) < \Phi(\underline{\sigma_B}) & \Rightarrow & \Psi(\underline{\sigma_A}) < \Psi(\underline{\sigma_B}) & \end{array} \tag{3.107}$$

Recalling that the two quadratic forms $\Phi(\sigma_{xx}, \sigma_{yy}, \tau_{xy})$ and $\Psi(\sigma_{xx}, \sigma_{yy}, \tau_{xy})$ implicitly represent two ellipsoids in a reference system $O\sigma_{xx}\sigma_{yy}\tau_{xy}$, it can be stated that, for a couple generic tension states $(\underline{\sigma_A}, \underline{\sigma_B})$, the statements (a), (b) and (c) are the only possible cases, if the two ellipsoids does not intersect in the $\sigma_{xx}\sigma_{yy}\tau_{xy}$ space.

It will be demonstrated now that this condition is fulfilled. Let us consider the reference system change defined by the matrix

$$\begin{aligned} &\underline{\underline{R}} = \tfrac{1}{2} \cdot \begin{vmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{vmatrix} \quad \Rightarrow \quad \underline{\underline{R}}^T \cdot \begin{vmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & c2 \end{vmatrix} \cdot \underline{\underline{R}} = \\ &= \frac{1}{4} \cdot \begin{vmatrix} \sqrt{2} & \sqrt{2} & 0 \\ -\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{vmatrix} \cdot \begin{vmatrix} a & b & 0 \\ b & a & 0 \\ 0 & 0 & c \end{vmatrix} \cdot \begin{vmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{vmatrix} = \\ &= \frac{1}{4} \cdot \begin{vmatrix} \sqrt{2}(a+b) & \sqrt{2}(a-b) & 0 \\ -\sqrt{2}(a-b) & \sqrt{2}(a+b) & 0 \\ 0 & 0 & 2c \end{vmatrix} \cdot \begin{vmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{vmatrix} = \\ &= \frac{1}{4} \cdot \begin{vmatrix} 4(a+b) & 0 & 0 \\ 0 & 4(a-b) & 0 \\ 0 & 0 & 4c \end{vmatrix} = \begin{vmatrix} a+b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & c \end{vmatrix} \end{aligned} \tag{3.108}$$

which represents a rotation of the reference system around the axis τ_{xy}. According to this change of reference system, it is possible to express the two ellipsoids:

$$\begin{aligned} &\tilde{\Phi}(\underline{x}) = \tilde{\Phi}(x, y, z) = \Phi(\sigma_{xx}, \sigma_{yy}, \tau_{xy}) \\ &where \quad \underline{x} = \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \end{vmatrix} \cdot \begin{vmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \tau_{xy} \end{vmatrix} = \underline{\underline{R}} \cdot \underline{\sigma} \implies \\ &\tilde{\Phi}(\underline{x}) = \underline{x}^T \cdot \underline{\underline{R}}^T \cdot \underline{\underline{D}}^{-1} \cdot \underline{\underline{R}} \cdot \underline{x} = \frac{1}{E} \cdot \underline{x}^T \cdot \begin{vmatrix} 1-\upsilon & 0 & 0 \\ 0 & 1-\upsilon & 0 \\ 0 & 0 & 2 \cdot (1+\upsilon) \end{vmatrix} \cdot \underline{x} \\ &\tilde{\Psi}(\underline{x}) = \tilde{\Psi}(x, y, z) = \Psi(\sigma_{xx}, \sigma_{yy}, \tau_{xy}) \quad where \quad \underline{x} = \underline{\underline{R}} \cdot \underline{\sigma} \implies \end{aligned} \tag{3.109}$$

$$\Longrightarrow \tilde{\Psi}(\underline{x}) = \underline{x}^T \cdot \underline{\underline{R}}^T \cdot \underline{\underline{D}}^{-1} \cdot \underline{\underline{R}} \cdot \underline{x} = \underline{x}^T \cdot \begin{vmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 3 \end{vmatrix} \cdot \underline{x}$$

The values of the elements of the two diagonal matrixes are the values of the three characteristic radii of the two ellipsoids: since v is a positive number (usually) it is easy to check that, for both the ellipsoids, the radius laying on the x axis is minor than the one laying on the y axis, which is minor than the radius laying on the z axis. Stated in an equivalent way, the three eigenvalues of the matrices $\underline{\underline{D}}^{-1}$ and $\underline{\underline{T}}$ have the same magnitude order ($\lambda_1 < \lambda_2 < \lambda_3$). It is easy to prove that this ensures that the two ellipsoids do not intersect each other, and, consequently, only the conditions (a), (b) and (c) are possible for the two quadratic forms.

3.7.3 Degenerate solutions for the resistance problems

At the very beginning of this chapter, it had been stated that the minimum compliance formulation is the easiest topology optimization problem, and that the resistance criteria are more difficult to implement. Such remark is not in contradiction with the considerations done in the previous section: in fact, it is important to highlight that what has been just demonstrated is that the sensitivity of a rejection criterion based on the strain energy produces the same effect of the one based on the equivalent Von Mises stress; the effect is limiting the average of both resistance and elastic energy, and doesn't say anything about wat appends at a local level.

An effective resistance constraint must be imposed at local level, in order to ensure not to reach high stress levels in any point of the structure. Ideally, it could be done by introducing a constraint to the optimization problem for every elements of the design space. This, for instance, would be a problem for the OC methods, which is not able to efficiently deal with a large number of constraints. To solve this problem, some techniques have been developed in order to aggregate the single element constraints [25].

Anyway, there is another problem to deal with, and it is the fact that, in presence of constraints that limit the equivalent stress in the structure elements, the optimal solution

exists in a degenerate solution space, and for this reason, the gradient based methods are not able to converge to a correct solution. Even in this case some relaxation techniques have been identified in order to allow the optimization algorithm to reach the ideal solutions. Nevertheless, a number of open issues are still open about this topic, and many researches are still active on this field. Let us consider the following example [48]: a structure is composed by two beams as are depicted in figure 3.6(a), characterized by the parameters reported in table 3.1.

	Lenght (L)	Density (ρ)	Young mod. (E)	max stress (σ^{max})	max stress (σ^{min})	Section (A)
Beam 1	1	1	1	1	-1	x_1
Beam 2	α	β	1	1	-1	x_2

Table 3.1: parameters for the 2 bars optimization problem

The equilibrium equations for the solution of the structure are:

$$\begin{cases} f_1 = -\dfrac{E_1 \cdot A_1}{L_1} \cdot \Delta u \\ f_2 = \dfrac{E_2 \cdot A_2}{L_2} \cdot \Delta u \\ f_1 + p = f_2 \end{cases} \quad \Longrightarrow \quad \begin{cases} f_1 = -x_1 \cdot \Delta u \\ f_2 = \dfrac{x_2}{\alpha} \cdot \Delta u \\ f_1 + p = f_2 \end{cases} \tag{3.110}$$

where Δu is the displacement of the point of application of the force p. Eliminating Δu from the expressions of f_1 and f_2:

$$\begin{cases} f_1 = -\dfrac{x_1 \cdot \alpha}{\alpha \cdot x_1 + x_2} \cdot p \\ f_2 = \dfrac{x_2 \cdot 1}{\alpha \cdot x_1 + x_2} \cdot p \end{cases} \quad \Longrightarrow \quad \begin{cases} \sigma_1 = -\dfrac{\alpha}{\alpha \cdot x_1 + x_2} \\ \sigma_2 = \dfrac{1}{\alpha \cdot x_1 + x_2} \end{cases} \quad if \;\; p = 1 \tag{3.111}$$

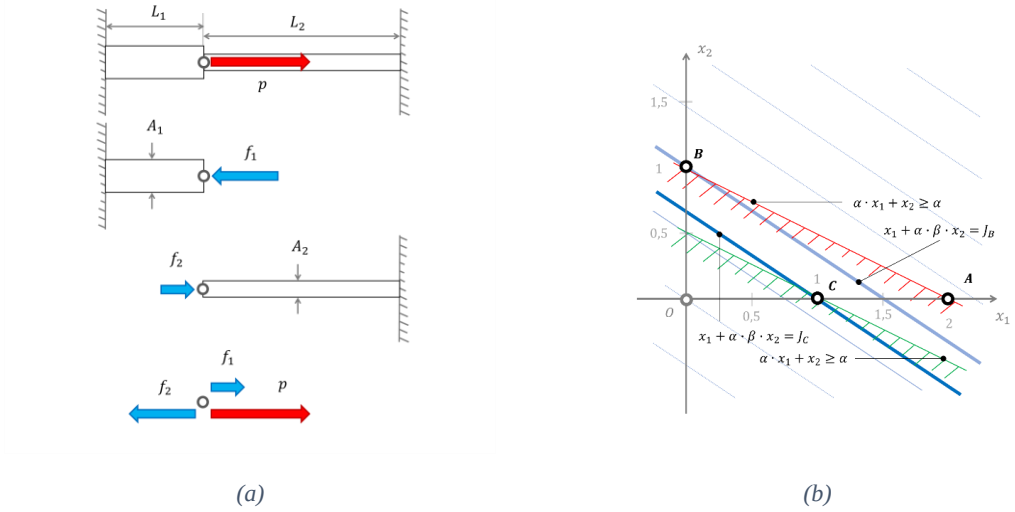
Where, as a further hypothesis, the module of the external load is set up to 1. Let us now consider the following optimization problem

$$\begin{aligned} min \quad & J(x_1, x_2) = x_1 + \alpha \cdot \beta \cdot x_2 \\ s.t. \quad & \sigma_1 = \frac{\alpha}{\alpha \cdot x_1 + x_2} \leq \left|\sigma_1^{min}\right| = |\sigma_1^{max}| = 1 \qquad if \quad x_1 > 0 \\ & \sigma_2 = \frac{1}{\alpha \cdot x_1 + x_2} \leq \left|\sigma_1^{min}\right| = |\sigma_1^{max}| = 1 \qquad if \quad x_2 > 0 \\ & x_1, x_2 \geq 0 \\ & x_1 > 0 \;\; and(or) \;\; x_2 > 0 \end{aligned} \tag{3.112}$$

which represents the minimization of the mass of the structure subject to resistance constraints; rearranging the inequality equations leads to

$$\begin{aligned} &min \quad && J(x_1, x_2) = x_1 + \alpha \cdot \beta \cdot x_2 \\ &s.t. \quad && \alpha \cdot x_1 + x_2 \geq \alpha && if \quad x_1 > 0 \\ & && \alpha \cdot x_1 + x_2 \geq 1 && if \quad x_2 > 0 \\ & && x_1, x_2 \geq 0 \\ & && x_1 > 0 \ \ and(or) \ \ x_2 > 0 \end{aligned} \tag{3.113}$$

which is a linear programming problem indeed. Figure 3.6(b) depicts the graphical interpretation of the problem: the clear grey lines are isocurves of the objective functions, and the red and green lines are the boundaries of the inequality constraints, if the values $\alpha = 0.5$ and $\beta = 3$ are adopted.



(a) *(b)*

Figure 3.6: set up for the 2 bars optimization problem (a) and diagram of the solutions (b)

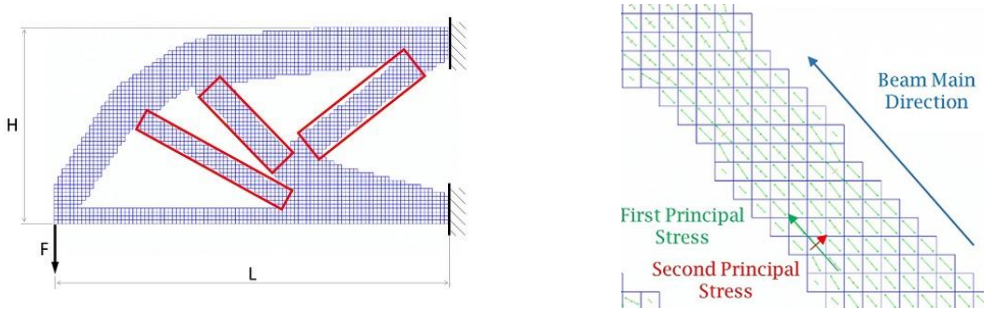
It can be noticed that, carrying out a gradient based solution, the algorithm would decrease the value of the objective function following its gradient, having the inequality constraints not active; once the point C is reached, the Lagrange multiplier of the constraint $\alpha \cdot x_1 + x_2 \geq \alpha$ is not longer null, ant the steps of the algorithm follow the path along the red line. Finally, it if fount a stationary point at the point B, with the value of the objective function equal to $J_B = 1.5$. Nevertheless, it is easy to verify that this is not the optimal solution, that actually may be fount in correspondence of the point C, being $J_C = 1$.

The fact that the gradient-based method is not able to reach the global solution depends by the fact that it belongs to a degenerate space of solutions, which corresponds to the segment $\overline{AC}$ [49].

3.8 Application of the directionality paradigm to topology optimization

In this section, according with the general framework introduced in chapter 1, the implementation of the directionality paradigm in an actual topology optimization algorithm will be presented. The following outcomes have been presented to the scientific community in occasion of an international conference, and have been further disclosed in a journal paper [50].

In order to show the main idea, figure 3.7 depicts a simple example of a cantilever truss subjected to a vertical force. Figure 3.7(a) reports the result of the optimization of a truss, obtained using the BESO algorithm, which was implemented by using the Python script for Abaqus described by Zuo and Xie [51].



(a) *(b)*

Figure 3.7: solution of a typical optimization procedure based on the BESO approach (a) and detail of a beam like feature (b)

A vertical force F is applied in the left bottom vertex of the truss, and the constraints neglect the displacements of the nodes on the right side. As it had been already highlighted, the result of a topology optimization is caried out using a BESO optimization strategy leads to a result highly compatible with the topology of the Mitchell trusses. For this reason, it does not surprise finding some sub-structures, which may be interpreted as beams, in the resulting topology, as it is shown in figure 3.7(a). Actually, the sub

structures are interpreted as beams because there is a single dominant principal direction, figure 3.7(b), or, in other words, they are subject to only traction or compression. From a structural point of view, this configuration is highly advantageous: in fact, it can be recalled that the Michell trusses are structures composed by elements subject to pure compression or traction, and, furthermore, are they show the lower product of volume and compliance.

According to the ontological framework presented in the previous chapter, it can be said that the truss is the structure, and the beams sub-structures are elements at the macro level. Furthermore, it is possible to say that, at a micro level, the beams elements are structures, and the isoperimetric square elements belonging to the original mesh are their constitutive elements indeed. In this sense, it is important analyse the passage from a larger dimensional scale to a smaller one. In fact, as a first important remark, it can be observed that the beam sub-structures are not disposed in the same direction of the square elements they are composed of, and this depends on a general initial definition of the mesh.

Restoring the consistency between the macroscopic result of the TO and the analysis of the state of tension of its elements may support the foundation of an original optimization methodology. This should be done according to the fact that the geometric features resulting by a topology optimization, which may be recognized as beams, are subjected to simply traction or compression, and not subjected to bending, for instance. This means that an enhanced distribution of material occurs inside the workspace if the elements are purely compressed or in traction.

3.8.1 Methodology

The proposed methodology has been devised in order to incorporate information about the stress configuration in the definition of the mesh. This is done by carrying out different stages, by the implementation of different modules: every step corresponds to a certain operation applied to the workspace, such as the structural analysis for the determination of stresses and strains, or the rearrangement of the model mesh. As a choice, it had been preferred to develop all the framework in a unique programming environment, MATLAB, instead of create a chain of different applications

communicating each other, using, for instance, a commercial software for FEM analysis, another application for the remeshing, and so on.

All the procedure is designed in order obtain a structure coherent with the empirical observations illustrated in figure 3.7, and, consequently, overcome the jagged appearance of the beams sub-structures. The whole process based on the following steps:

- discretization and mesh generation;
- rough BESO optimization: this procedure comprises a loop which involves the following, standard steps:
    - evaluation of the strain energy of each element by the mean of finite element analysis;
    - sensitivity analysis;
    - application of a convolution filter;
    - update of element densities according and evaluation off the convergence conditions;
- stress configuration evaluation, and determination of the principal directions: this step is based on the result of the last finite element analysis of the previous stage;
- element rearrangement: the re-meshing comprises the rotation and connection based on the tension flux within the structure material. To perform this stage, a routine for the rearrangement of the geometry has been developed.;
- size optimization.

In the following subsections, each step will be described.

Discretization

This initial stage is common to all topology optimization procedures, and it comprises the definition of a feasible workspace which contains the continuum structure, and the division of this workspace in a set of elements. In this case, the mesh is composed by all equals square isoparametric elements. Adopting this type of discretization, it is possible compute stress and strains in the centre of mass of the elements, by averaging the mean quantities in correspondence of the Gauss points. Furthermore, once the nodal displacements are determined, it is possible to calculate the strain energy of a single element by the use of the quadratic form

$$E_{ele} = \frac{1}{2} \cdot \left(\underline{U}_{ele}\right)^T \cdot \underline{\underline{K}}_{ele} \cdot \underline{U}_{ele} \tag{3.114}$$

being $\underline{U_{ele}}$ the vector of the nodal displacements of the elements, and $\underline{\underline{K_{ele}}}$ the stiffness matrix at element level.

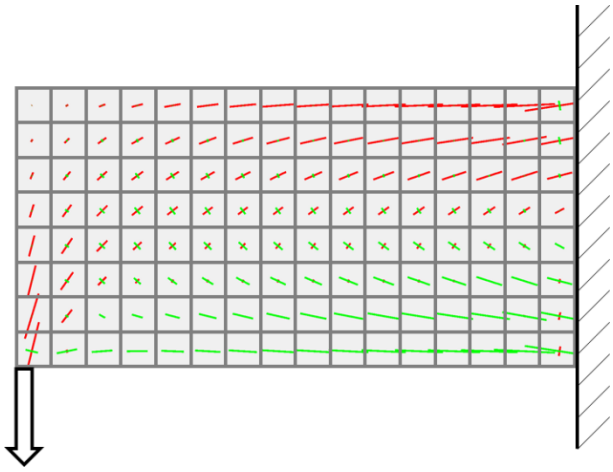


Figure 3.8: design space discretization using a mesh of isoparametric square elements

Figure 3.8 depicts an example of discretization phase, and, moreover as an anticipation of the next step, it shows the results of the initial FE analysis, highlighting the principal stresses slope and modulus, σ_I and σ_{II}, in correspondence of the elements. These quantities are already expressed in the principal, local system of reference, and are derived by the components of the stress tensor σ_x, σ_y, and τ_{xy}, expressed in the original reference system of the element.

Rough BESO optimization

The second phase consists in an initial rough Topology Optimization, based on the output initial mesh defined in the previous step. The compliance minimization problem is considered in here, and it is subject to a volume constraint. According to the previous sections, let us recall that the mathematical formulation for the BESO topology optimization problem is given by the system (3.91)

$$
\begin{aligned}
min \quad & J\left(\underline{\rho}\right) = \frac{1}{2}\left(\underline{u}\left(\underline{\rho}\right)\right)^{T} \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) \\
s.t. \quad & \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}\left(\underline{\rho}\right) = \underline{f} \\
& V\left(\underline{\rho}\right) - V_0 = 0 \\
& \rho_i = \begin{cases} 1 \\ \rho_{min} \end{cases} \quad i = 1, \dots, n
\end{aligned}
$$

Both hard kill and soft kill BESO methods may be adopted in order to carry out the optimization process: here the soft kill version is adopted. The solution algorithm is well known from literature [52], and it is depicted in figure 3.9(a).

Step (1): once all initial parameters of the optimization are set up (volume fraction V_0, initial densities of the elements ρ_i, etc…), a first FE analysis is caried out: usually the process starts with the design domain full of material ($\rho_i = 1\ ;\ \forall\ i$), and this means that the young modulus of every element is such that $E = E(\rho_i) = E_0$, and, consequently, $\underline{\underline{D}}(\rho_i) = \underline{\underline{D}}_0$. This means that the initial global stiffness matrix $\underline{\underline{K}}_0$ can be computed, and so it is possible to determine the initial vector of nodal displacements $\underline{u}^{(0)}$:

$$
\underline{u}^{(0)} = \left(\underline{\underline{K}}_0\right)^{-1} \cdot \underline{f} \tag{3.115}
$$

Step (2): BESO is a gradient based method, so it requires the computation of the sensitivities of the objective function $f\left(\underline{\rho}\right)$ with to respect the design variables ρ_i, and it can be done applying the adjoint method reported in appendix A.

Step (3): once the sensitivities are computed, usually they are modified by the application of a blurring filter that actually average the value of the sensitivity of every element with the sensitivity of the neighbour elements. The goal of this procedure is overcome the phenomenon of the mesh dependency of the result of the optimization process: this effect consists in the fact that, discretizing the workspace by the means of smaller and smaller elements (consequently increasing the number of elements), the result of the topology optimization does not converge to a well-defined shape, but, on the contrary, has the tendency to create microstructures, or, as it is reported in many researches, checkerboard patterns. The application of a convolution filter actually imposes a limit to the dimension of the structures at a microscopic level, that are equivalent, according to the proposed ontological framework, to the dimension of the elements at a macroscopic level.

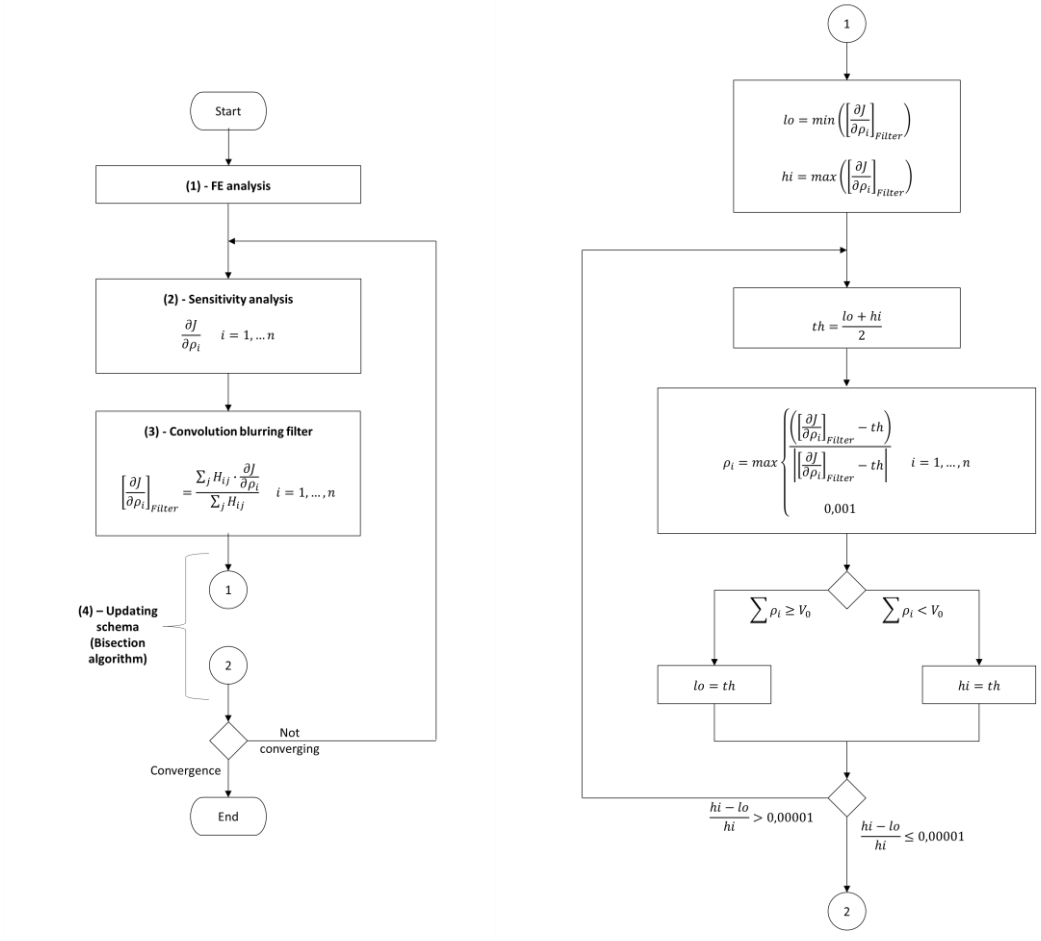
The averaged sensitivity of the element i is

$$\left[\frac{\partial f}{\partial \rho_i}\right]_{Filter} = \frac{\sum_j H_{ij} \cdot \frac{\partial f}{\partial \rho_i}}{\sum_j H_{ij}} \quad i = 1, \dots, n \tag{3.116}$$

where the index j corresponds to the set of neighbours of the element i according to $dist(j,i) \leq R$, being the distance between the centres of the j-th element and i-ts element; operator H_{ij} is defined by the following

$$H_{ij} = R - dist(j,i) \tag{3.117}$$

where R is the minimal characteristic dimension of the element at macro level and a structure at micro level.



(a) *(b)*

Figure 3.9: flowchart of the BESO optimization algorithm (a), and flowchart of the bisection routine (b)

Step (4): once the modified sensitivities are computed, it can be executed the actual updating process of the densities of the elements. The flux diagram of the procedure is shown in figure 3.9(b): it is basically the representation of a bisection algorithm, in which the density of the i-th element, ρ_i, is eventually set to the values 1 or 0.001, depending on the value of the correspondent sensitivity $\left[\frac{\partial J}{\partial \rho_i}\right]_{Filter}$; the elements densities are updated until the volume ratio is respected, and all the sensitivities of the material elements are greater than their average value.

The result of the TO consists of the definition set of elements with high strain elastic energy. Low energy elements are firstly flagged, and a low density, and, consequently, a low Young modulus, is assigned. In a second phase, the elements with low density are eliminated.

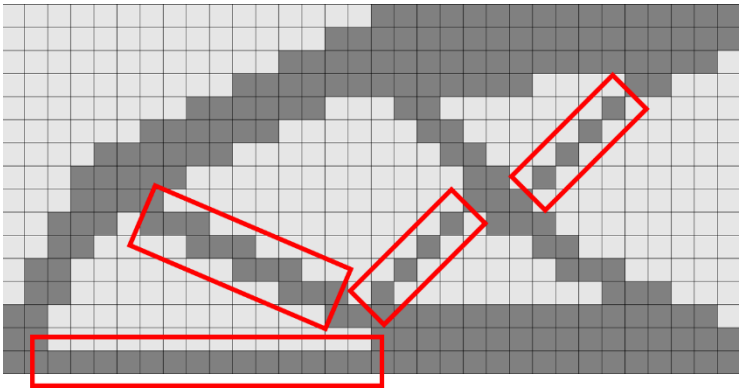


Figure 3.10: result of the rough BESO optimization process

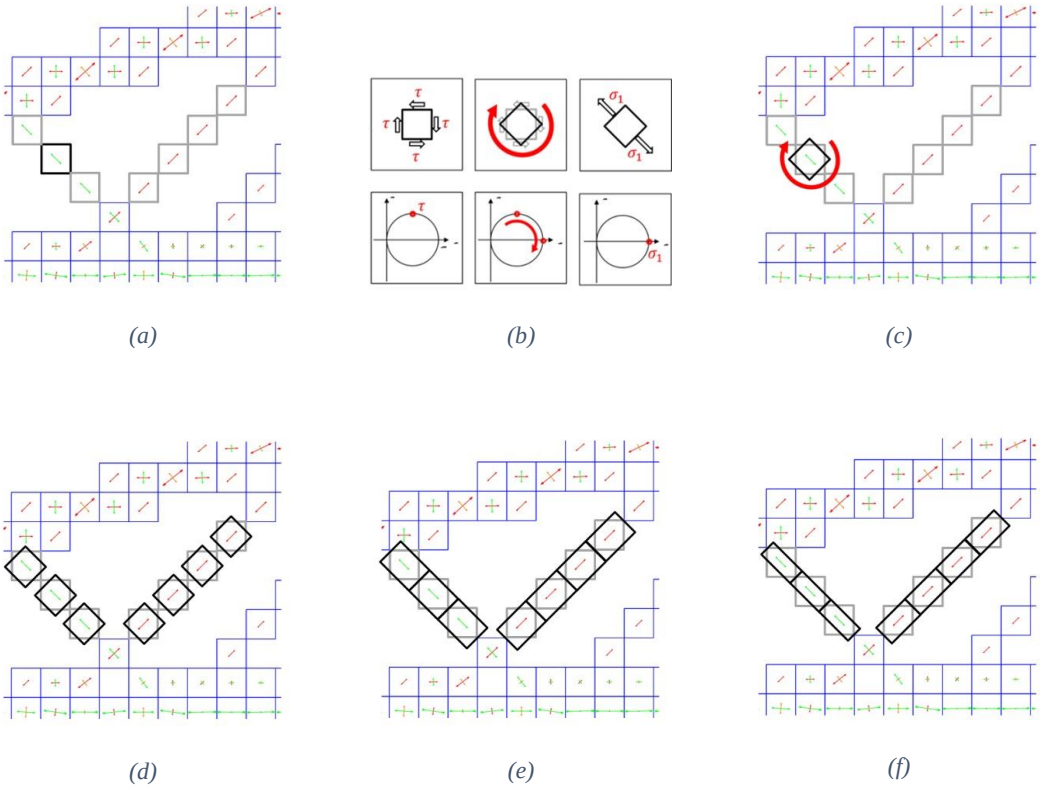
Figure 3.10 depicts an example of result of the rough BESO optimization; in the same figure, it is possible to identify the sub-structures of the beams, and some examples are highlighted. While investigating the nature of the stresses in such sub-structures, it is highlighted that the major principal tensions are oriented in the same direction of the elements as indicated by the red rectangles.

Stress configuration evaluation

After the set of elements with high strain energy have been identified (in figure 3.10 are dark-grey elements), the next step is the elimination of the elements characterized by a low level of strain energy (light-gray): this is done modifying the connectivity matrix of the mesh, erasing the lines in correspondence of void elements.

In fact, the outcome of the topology optimization process is the determination of the actual material distribution of the structure in terms of elements which are in correspondence of existing material, and elements occupying a portion of space where the material does not exist. Differently by the SIMP method, in BESO the elements have only two feasible densities, $\rho_i = 1$ or $\rho_i = \rho_{\min} = 0.001$: for this reason, it is possible to distinguish the elements in which the computation of the stresses and strains have a physical meaning, elements corresponding to material, and elements in which such computation is not significant. As a remark, it can be recalled that, even if the element densities updating is different, BESO adopt the SIMP material interpolation schema.

Once the mesh composed by only material elements is defined, a new finite element analysis of all the structure is performed, and each element is considered. A MATLAB function calculates the principal stresses and strains, and the correspondent slope for the principal directions.



(a) *(b)* *(c)*

(d) *(e)* *(f)*

Figure 3.11: re-orienting, connection and size optimization of the mesh elements

Therefore, the resulting structure of the topology optimization a region of the 2D or 3D Euclidean space defined by the union of the material elements, and characterized by a jagged profile (2D) or surface (3D); furthermore, it is possible to consider the stress state only in the elements actually representing the continuum structure, and, based on the stress tensor, it is possible to identify the principal directions of the stress field, evaluated in the centre of the elements. The two principal directions of the i-th element, perpendicular each other, are identified by the mean of an angle $\theta^{(i)}$:

$$\theta^{(i)} = \frac{1}{2} arctg\left(\frac{2{\tau_{xy}}^{(i)}}{{\sigma_x}^{(i)} + {\sigma_y}^{(i)}}\right) \tag{3.118}$$

which is he double of the angle identifying the tension state of the element in the Mohr circle; furthermore, the angle $\theta^{(i)}$ represent the entity of the rigid body rotation required fort the transformation from the original local system, in which ${\sigma_x}^{(i)}$, ${\sigma_y}^{(i)}$, and ${\tau_{xy}}^{(i)}$ are expressed, to the principal local system, in which ${\sigma_I}^{(i)}$ and ${\sigma_{II}}^{(i)}$are expressed. This operation is shown in the figures 3.11(a), (b) and (c).

Elements rearrangement

The further stage of the proposed methodology is the rearrangement of the elements of the structure mesh: this is based on the operation of rigid rotation according to the slope of the principal stress tensor, and it is carried out in order to recover the continuity of the mesh, obtaining, consequently, a smoother perimeter. In other words, the goal is considering the new information regarding the orientation of the principal stress tensor directions to create a new mesh.

As depicted in figure 3.11(d), all the elements are rotated coherently to the slope of their principal system of reference. Obviously, due to such rotation, it is not possible to preserve the continuity of the material. For this reason, the geometry of the mesh is recovered. This is done joining the adjacent elements by sharing the corresponding nodes, as shown in figure 3.11(e).

The rotation process of the elements means that the decoupling of the nodes is performed. The initial mesh is a standard square mesh and each node is shared by four elements. Usually, the position of the nodes and the list of the nodes belonging to each element are stored in arrays. The rotation of the elements means that the vertices of the elements do not coincide anymore. For this reason, the vector of the nodal positions, and the vector of

the elements connectivity change, and, more precisely, while the dimensions of the connectivity matrix remains the same, the length of the array of the nodal positions changes. Furthermore, providing again the continuity means that adjacent elements will share the nodes after the rotation. Again, this is done modifying the nodes and elements arrays, decreasing this time the dimension of the nodes array.

Size optimization

Size optimization of the structures at micro level is the further step of the methodology: this operation is represented by the passage form figure 3.11(e) to figure 3.11(f). The convenience of this further shape/size optimization process is due to the rearrangement of the elements of the mesh. In fact, the elastic strain energy stored in every rotated element will be different from the value corresponding to the last step of optimization.

The main idea is that, in presence of the beam-type elements at macroscopic level, which are structures composed by the rearranged elements at a microscopic level, the resistance section depends only on the module of the first, and unique principal stress. The goal of such procedures is the definition of the size of the elements so that their elastic strain energy density is the optimal for every element.

After the re-definition of the mesh, the elastic energy stored may be written for the mono axial state of tension of the elements:

$$C = W_{Opt} = \mathrm{V}(\mathrm{s}_{Opt}) \cdot \frac{1}{2}\sigma\varepsilon \tag{3.119}$$

where S is the thickness of the element of the beam feature, and V is the volume (area) of the element itself.

Because some elements disclose a mono axial state of tension, it is possible to (locally) formulate the optimization problem using as optimization variables the geometric quantities is as follows:

$$\begin{aligned} min \qquad & C = V(\mathrm{s}_i) \cdot \frac{1}{2}\sigma\varepsilon \\ & s_i = (0, s_{opt}) \qquad\qquad i = 1, \dots, \bar{n} \end{aligned} \tag{3.120}$$

where $\bar{n}$ is the number of the solid elements of the optimized structure.

3.8.2 Results

The proposed method has been implemented in the MATLAB programming environment, without the use of any external package: the code has been written using the basic features of the development environment for the creation of both basic and graphical functions.

Figure3.12(a) and 3.12(b) show the averaged state of tension for every element of the full material structure: the former visualizes the three components of the (plane) stress tensor $\sigma_x{}^{(i)}$(blue vectors), $\sigma_y{}^{(i)}$(magenta vectors) and $\tau_{xy}{}^{(i)}$(green vectors), with respect to the original reference system, and the latter represents the principal tensions $\sigma_I{}^{(i)}$(green vectors) and $\sigma_{II}{}^{(i)}$(red vectors), and the rotated elements. The different grades of grey represent the different levels of strain energy of the elements (dark for higher strain energy, lighter for lower strain energy).

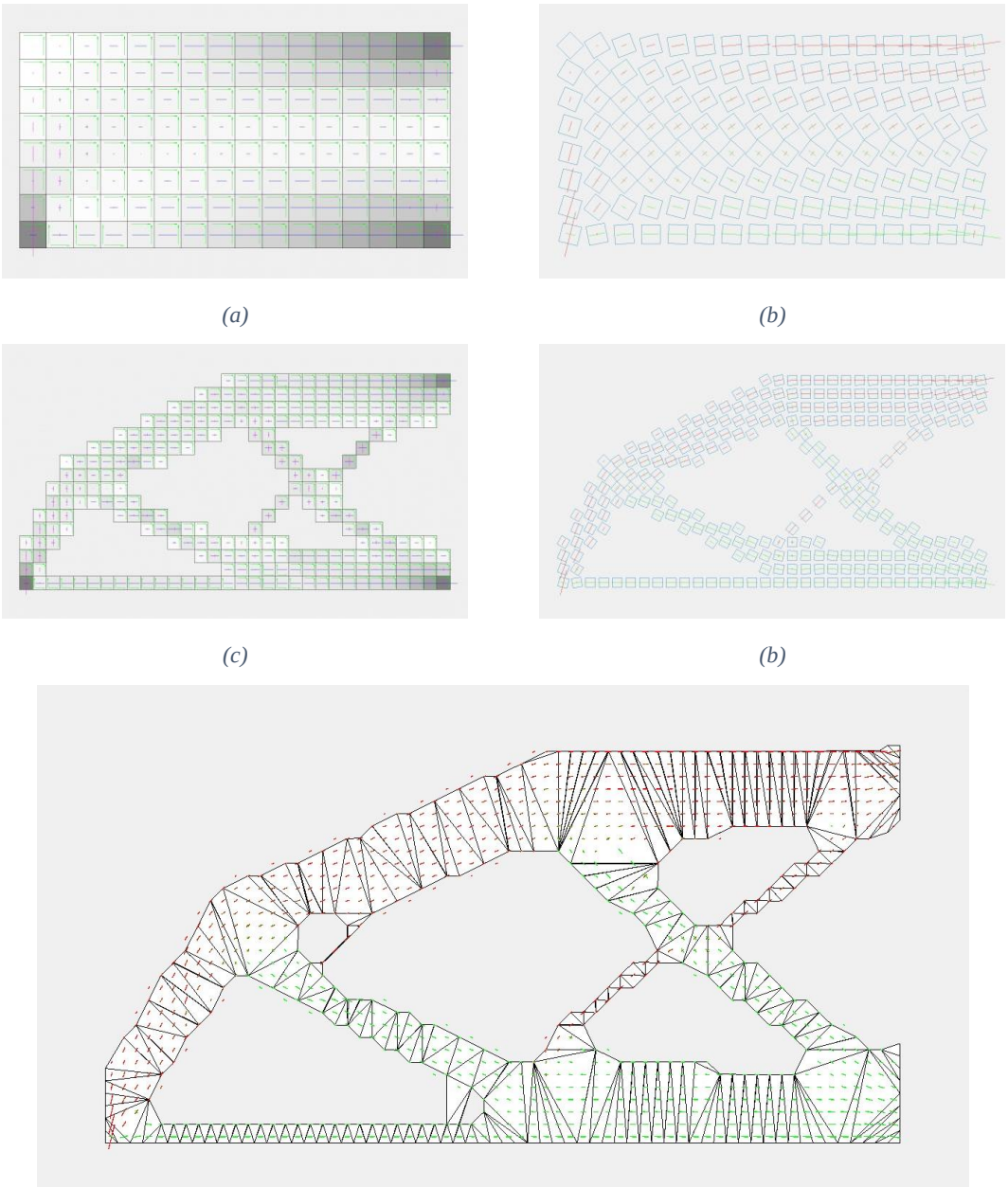
Furthermore, figure 3.12(c) and 3.12(d) depict the components of the stress tensor in the original and principal reference system for the structure obtained by the use of the soft kill BESO topology optimization process: this is the base for the further elaboration described previously, based on the rearrangement of the rotated elements, and the redefinition of the set of nodes and the redefinition of the incidence matrix. The result of all the process is reported in figure 3.12(e), and it may be noticed that is that the new geometry is less affected by the checkerboard appearance than the one shown in figure 3.12(c).

Comparison between BESO and proposed approach

A confront between the outcome of the standard soft kill BESO topology optimization procedure and the result of the proposed method has been caried out in order to evaluate the effectiveness of the implementation of the directionality in the topology optimization process.

The test has been done considering the boundary conditions specified in figure 3.7(a) (H=100mm, L=200mm, F=1000N). Different runs of the code have been done adopting different element sizes, and, more precisely, the design space has been initially discretized in a grid of 16x8 elements, then 20x10 elements, etc., as reported in figure 3.13.

Two indexes have been defined for the confront: the first one, $\%\, Area\ decrease$, represents the improvement of the proposed method in terms of area of the resulting structure; furthermore, the second index, $\%\, Compliance\ decrease$, represent the compliance decrease when the new approach is adopted.



(a) *(b)*

(c) *(b)*

(e)

Figure 3.12: example of graphical output of the MATLAB implementation of the proposed approach.

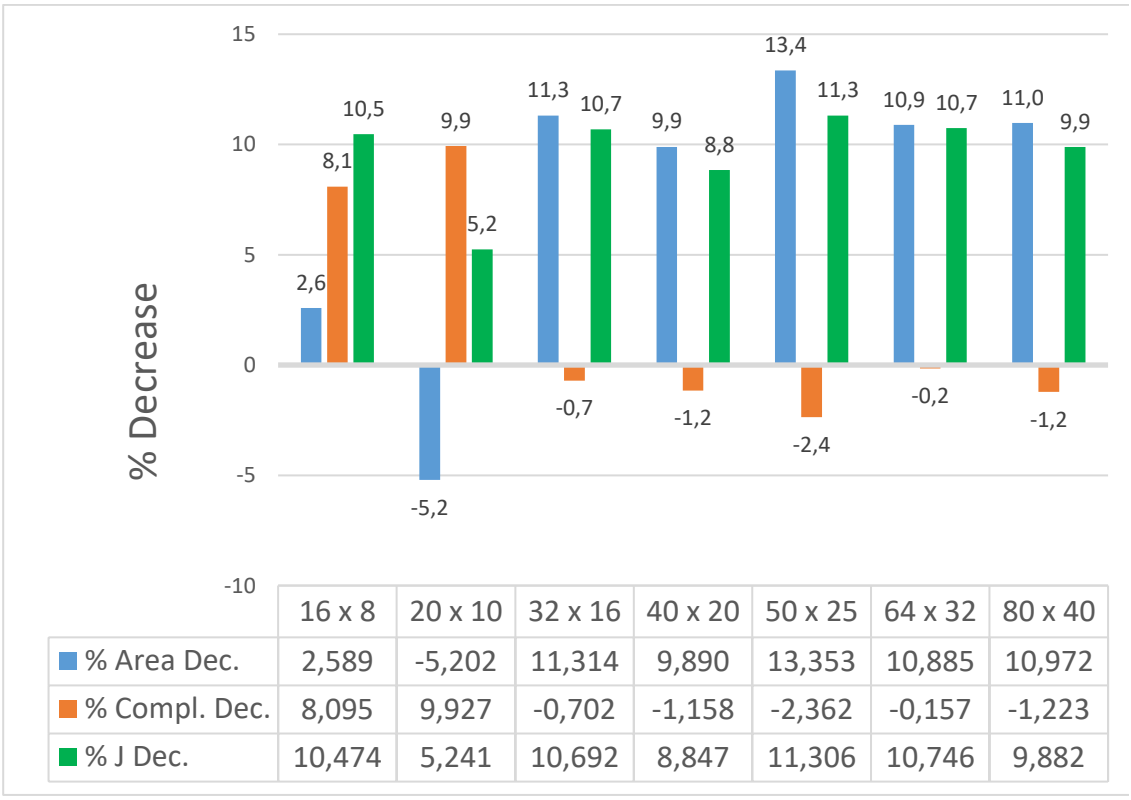
The expressions of the two indexes are the following

$$\%\ Area\ decrease = \frac{(\text{Area}_{BESO} - \text{Area}_{Re-Mesh})}{\text{Area}_{BESO}} \tag{3.121}$$

$$\%\ Compliance\ decrease = \frac{(\text{Compliance}_{BESO} - \text{Compliance}_{Re-Mesh})}{\text{Compliance}_{BESO}} \tag{3.122}$$

where Area_{BESO} and $\text{Area}_{Re-Mesh}$ are the areas of the structures produced by the soft kill BESO approach, and the proposed approach respectively; similarly, Compliance_{BESO} and $\text{Compliance}_{Re-Mesh}$ are the resulting compliances for the standard soft kill BESO method, and the modified one. The results of the comparison for different elements sizes are depicted in figure 3.13.

It can be noticed that most of the results are positive, which means that the proposed methodology gives better results. When a value is negative, for example we have an increase of the area, on the other hand we have a higher decrease (in percent) of the compliance, which means that globally there is a more efficient use of material.



	16 x 8	20 x 10	32 x 16	40 x 20	50 x 25	64 x 32	80 x 40
% Area Dec.	2,589	-5,202	11,314	9,890	13,353	10,885	10,972
% Compl. Dec.	8,095	9,927	-0,702	-1,158	-2,362	-0,157	-1,223
% J Dec.	10,474	5,241	10,692	8,847	11,306	10,746	9,882

Figure 3.13: results of the comparison between the BESO TO, and the proposed methodology.

Furthermore, figure 3.13 shows that the structure obtained by the use of the proposed method better satisfies the objective the formulation of the objective function for the BESO approach, derived by the formulation given by Mitchel, $J = W \cdot V$, and, for this reason, is a better approximation of the minimum cost truss.

A last remark regards the increase of the execution time introduced by the re-meshing procedure: table 3.2 reports the comparison of the computational cost of the standard BESO procedure, and the computational cost of the routine for the new discretization of the structure. It should be highlighted that the BESO procedure is a well-established procedure, implemented in a very efficient code; on the contrary, the implementation of the proposed methodology is still object of improvement, and far to be already optimized.

	16x8	20x10	32x16	40x20	50x25	64x32	80x40
BESO TO (sec)	1.3199	1.3536	1.4465	1.4509	1.5987	2.2967	2.5848
RE-Mesh (sec)	25.7587	21.6436	24.0930	25.7886	31.3968	41.2450	69.6368

Table 3.2: comparison of the computational cost of BESO and the implementation of the proposed approach

Application of the proposed approach to tridimensional problems

The previous analysis focused on the application of the proposed methodology to the definition of a 2D structure; nevertheless, since tridimensional problems are more relevant for technical applications, some considerations must be addressed on the extension of the procedure to the synthesis of structures in space.

The remeshing process can be easily extended to a hexahedral discretization of a 3D design space, and the described algorithm may be integrated with minimal conceptual changes. Unfortunately, one of the main issues in the passage from 2D to 3D TO, is the increase of the computational cost.

For instance, even if BESO, as reported in table 3.2, shows good performances in bi-dimensional problems, in 3D TO the execution time rises dramatically even adopting coarse hexahedral meshes: for this reason, even if a finer mesh is required in order to identify in a proper way all the geometrical features of the optimized solution, the designer must find a trade-off between the quality of the result, and computational cost.

Another important aspect is the filtering of the values of the design variables and sensitivities: in fact, increasing the number of elements discretizing the design space,

implicates that it is necessary to apply a convolution filtering, setting an adequate radius in order to obtain feasible geometries. Again, in the case of 2D applications, the computational cost keeps low even with medium fine meshes, but rises dramatically in 3D applications.

Consequently, keeping the discretization mesh as coarse as possible should be the way to limit computation time in 3D applications. This is true as long as TO provides an indication of the ideal topology of the structure, subject to a particular load configuration: if the geometrical features are identified, it is always possible to model a posteriori the actual structure, eventually carrying out a second stage of shape or size optimization. Nevertheless, keeping a coarse discretization of the design space requires a regular behaviour of the solution with respect to the refinement of the mesh, but, as it had already been addressed, continuum TO is affected by the checkerboard issue, and requires filtering: this means a finer discretization of the workspace, and a rise of the computation time.

This is the scenario while adopting a regular mesh as the one depicted in figure 3.8, but what if we should be able to build the mesh with elements oriented according to the directions of the principal stresses? Some preliminary tests suggest that the proposed re-meshing strategy should allow to obtain a well-defined topology even adopting a coarse mesh, compensating the increase of the computational cost with respect to the simple BESO procedure, with the decrease of elements number.

Indeed, all these aspects must be better investigated: in the proposed methodology, the remeshing process is applied at the end of the optimization procedure, and allows a refinement of the solution without affecting much the resulting topology. Nevertheless, the next stage of the research is involving the modify of the discretization of the design space during the optimization process itself, according to what will be introduced in the next section.

This leads to a last remark about the extension of the remeshing procedure to SIMP: in the previous section, the choice of BESO as optimization procedure is justified by the binary nature of the result, which allows to obtain the final geometry by the rotation of the full material elements, and recovery their connectivity. On the contrary, results generated by SIMP exhibit elements with intermediate densities: this means that the re-

meshing procedure should involve a larger part of the deign space, or all the design space, which is expected to be computational costly, but applicable in principle.

3.9 Extension of the proposed methodology

In the previous sections, it had been described a strategy to determinate how to reorient and rearrange the finite elements in order to modify the mesh after a first step of topology optimization. Anyway, as it had been highlighted in the previous subsection, even if the re-mesh procedure is suitable to refine the results, it doesn't affect much the resulting topology compared standard optimization approaches.

A better strategy would be the one which allows the designer to figure it out how the material would naturally dispose in order to better react to a stress field caused by a particular load configuration: to do this, it is possible to better implement the directionality biomimetic paradigm.

Figure 3.14 recalls that many researches showed how the fibres of a bone naturally dispose according to the force lines of the stress field [53]. The observation of such structure reveals that there is a web of fibres which contribute to the structural integrity of the tissue: this is in concordance with the nature of a generic planar stress state, because, in every point two principal stresses may be identified, perpendicular one another. This architecture is the equivalent of the Michell trusses Hencky net, obtained by a generative process in which the bone dynamically distribute material according to the changes of load conditions.

More in detail, the features in said diagram are:

- red lines represent the flux lines for the first principal direction (σ_I);
- green lines represent the flux lines for the second principal direction (σ_{II}).

Beside these physical quantities, it would be convenient consider the elastic (distortion) energy W_{dist}, which is a well-known parameter: for every element, the angle θ has been already defined, such as the modules σ_I and σ_{II}. Consequently, the local principal system of reference, have been already computed, it is possible to obtain the elastic (distortion) energy stored in the element:

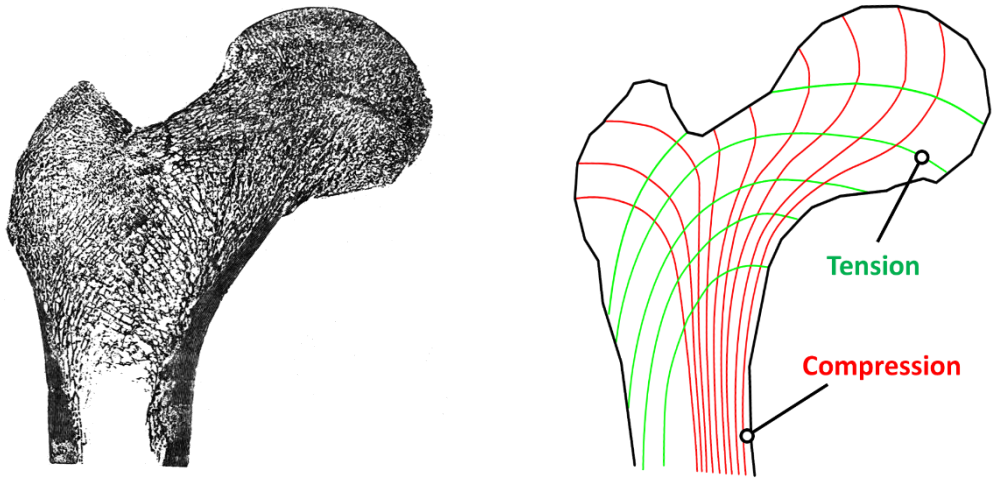
$$W_{dist} = \sqrt{\sigma_I^2 - \sigma_I \sigma_{II} + \sigma_{II}^2} \tag{3.123}$$

The main idea is the using the information regarding the stress distribution in the work-space in order to predict a priori the ideal material distribution.

More in depth, two fields are identified, the one comprising the first principal stress σ_I information, and the second comprising the second principal stress σ_{II} information. It is possible to draw the stream lines of both fields: one of the properties is that such lines are mutual perpendicular.

The next stage proposed for the creation of a new topology optimization tool starts from the mapping of the tensor stress in the continuum. In order to obtain a structure with the same features as the bone tissue, a procedure is proposed which is divided in three stages:

1) calculate the tensor field representing the tension state of the structure not yet optimized;
2) determine the topology of the tensor field;
3) according to the topology of the tensor field, the model of the optimized structure is generated: the model represents the optimized domain according to the objective functions. Moreover, the geometry is defined using a mesh which is natively suitable for the isogeometric analysis.



(a) *(b)*

Figure 3.14: lines of principal stress in femoral bone

3.9.1 Determination of the tensor field

The first step of the previous schema is carried out providing the solution of the linear elastic analysis of a structure which is subject to a system of forces and boundary conditions, using the finite element method. Even if this step has been already implemented, and actually the test code is already provided of the necessary the FEM routines at the present days, nevertheless, the development of the finite element procedure should be improved. As an example, isogeometric analysis [54] could be taken in account in order to provide a model that better describes the state of tension according to the geometric features generated by the optimization process, such as beams like sub structures. A similar idea has been reported in the journal paper based on the present research [50], where the possible use of NURBS has been proposed, in order to recover a smooth geometry for the structure perimeter.

3.9.2 Definition of the topology of the tensor field

The study of topology-based methods in visualization is a field of research which has been largely investigated in the past years. One of its sub-topics is the tensor field visualization: in analogy to the visualization of a vector field, a symmetric second order tensor field may be visualized as a web of always perpendicular streamlines. The tensor streamlines are tangent to the principal directions of the tensor in every point of the domain, determined as the direction of its eigenvectors. As had been already stated, an ideal disposition of material is coherent to the principal directions for the tensor of tensions. This outcome comes from both results of topology optimization and biomimetic observations.

As direct consequence of the reasoning, the proposal for the further stage of the research is develop an algorithm which defines the optimized geometry for the linear elastic structure starting from the topology of the symmetric second order tensor representing

the tensions field of the structure, or, in other words, the web of streamlines for the principal directions.

Many researchers investigated such problems [55][56] and procedures have been developed in order to define the tensor topology in 2D and 3D domains.

3.9.3 Optimized model generation

Once the web of one-dimensional entities (tensor field streamlines), this may be used as ruler for the generation of the 3D actual geometry. This may be done in many different ways, but, basically, in order to avoid the well-known drawbacks of the use of a polygonal mesh, an alternative is the use of a NURBS-based definition of the geometry, which is at the base of isogeometric analysis.

Even if a further literature research is necessary, it is reasonable thinking that a domain which is defined on the basis of geometries which incorporate the information about the stress configuration of the structure itself would show an ideal behaviour in terms of structural response.

3.10 Conclusions and future developments

Chapter 3 has been devoted to the investigation of the possibility of implementing the biomimetic directionality paradigm in design tools; as it had been widely documented in literature, topology optimization (TO) codes are one of the most promising possibilities to generate structure geometries in an automated way.

TO is usually based on finite element method (FEM), and in most application the design domain is discretized with a regular a priori defined mesh; this implies that, for replicating the theoretical results of fully stressed structures described in Mitchell's work, it is necessary using a high number of elements, in order to mitigate the jagged aspect of the beam-like features.

To overcome these undesired drawbacks, it had been here proposed a modified procedure involving the reconstruction of the mesh according to the slope of the principal directions of the stress tensor.

More in detail, the first part of the chapter provided a report about the principal elements for the construction of a topology optimization algorithm: the finite element analysis, the material interpolation schemes and gradient based constrained optimization are all the topics reviewed in order to outline the necessary theoretical base. Moreover, two popular approaches, SIMP and BESO, have been analysed and compared, and it had been described the importance of which objective function is taken in account in order to carry out the optimization process. Finally, the main limitations of the existing methodologies have been highlighted, and it had been addressed which are the open issues in the field.

The second part of the chapter deals with the proposal of a modified BESO approach, which ideally should be taken in account the configuration of the stress field, and in particular the slope of the principal directions of the tensor field. Based on this analysis, a procedure is proposed which includes the definition of a new mesh able to better represent the sub-structures that naturally rise in the optimization process. This procedure potentially offers different advantages: firstly, a design that better representation of the feature of the structure is obtained, providing a better base for the creation of the final design. Moreover, the bean like substructures can be subject to a further size optimization, and this may mitigate the difficult of the optimization procedure to ensure the fulfilment of the resistance criteria, which is one of the issues highlighted by the literature analysis.

The proposed topology optimization procedure has been implemented in a prototype MATLAB code, and the first result are encouraging. Nevertheless, further experiments must be caried out, and the procedure itself may be improved by a further development of the research.

3.11 Bibliography

[25] Bendsoe, M. P., & Sigmund, O. (2013). Topology optimization: theory, methods, and applications. Springer Science & Business Media.

[26] Cook, R. D. (2007). Concepts and applications of finite element analysis. John wiley & sons.

[27] Khennane, A. (2013). Introduction to finite element analysis using MATLAB® and abaqus. CRC Press

[28] Liu, K., & Tovar, A. (2014). An efficient 3D topology optimization code written in Matlab. Structural and Multidisciplinary Optimization, 50(6), 1175-1196.

[29] Saxena, A., & Sahay, B. (2007). Computer aided engineering design. Springer Science & Business Media.

[30] Nocedal, J., & Wright, S. (2006). Numerical optimization. Springer Science & Business Media.

[31] Fleury, C. (1989). CONLIN: an efficient dual optimizer based on convex approximation concepts. Structural optimization, 1(2), 81-89.

[32] Svanberg, K. (1987). The method of moving asymptotes—a new method for structural optimization. International journal for numerical methods in engineering, 24(2), 359-373.

[33] Svanberg, K. (2007). MMA and GCMMA-two methods for nonlinear optimization. vol, 1, 1-15.

[34] Andreassen, E., Clausen, A., Schevenels, M., Lazarov, B. S., & Sigmund, O. (2011). Efficient topology optimization in MATLAB using 88 lines of code. Structural and Multidisciplinary Optimization, 43(1), 1-16.

[35] Ananiev, S. (2005). On equivalence between optimality criteria and projected gradient methods with application to topology optimization problem. Multibody System Dynamics, 13(1), 25-38.

[36] Sigmund, O. (1994). Design of material structures using topology optimization (Doctoral dissertation, Technical University of Denmark).

[37] Xie, Y. M., & Steven, G. P. (1993). A simple evolutionary procedure for structural optimization. Computers & structures, 49(5), 885-896.

[38] Tanskanen, P. (2002). The evolutionary structural optimization method: theoretical aspects. Computer methods in applied mechanics and engineering, 191(47-48), 5485-5498.

[39] Michell, A. G. M. (1904). LVIII. The limits of economy of material in frame-structures. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 8(47), 589-597.

[40] Maxwell, J. C. (1870). I.—on reciprocal figures, frames, and diagrams of forces. Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 26(1), 1-40.

[41] Bouchitté, G., Gangbo, W., & Seppecher, P. (2008). Michell trusses and lines of principal action. Mathematical Models and Methods in Applied Sciences, 18(09), 1571-1603.

[42] McMillan, C., & Buffa, E. (1970). Mathematical programming : an introduction to the design and application of optimal decision machines.

[43] Huang, X., & Xie, Y. M. (2010). A further review of ESO type methods for topology optimization. Structural and Multidisciplinary Optimization, 41(5), 671-683.

[44] Nesterov, Y. (2003). Introductory lectures on convex optimization: A basic course (Vol. 87). Springer Science & Business Media.

[45] Boyd, S., Boyd, S. P., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press.

[46] Svanberg, K. (1994). On the convexity and concavity of compliances. Structural optimization, 7(1-2), 42-46.

[47] Li, Q., Steven, G. P., & Xie, Y. M. (1999). On equivalence between stress criterion and stiffness criterion in evolutionary structural optimization. Structural optimization, 18(1), 67-73.…

[48] Stolpe, M. (2003). On models and methods for global optimization of structural topology (Doctoral dissertation, Matematik).

[49] Rozvany, G. I. N., & Birker, T. (1994). On singular topologies in exact layout optimization. Structural optimization, 8(4), 228-235.

[50] Caputi, A., Weiss Cohen, M., Russo, D., & Rizzi, C. (2018). Topology Optimization using Explicit Stress Tensor Analysis.

[51] Zuo, Z. H., & Xie, Y. M. (2015). A simple and compact Python code for complex 3D topology optimization. Advances in Engineering Software, 85, 1-11.

[52] Huang, X., & Xie, Y. M. (2010). A further review of ESO type methods for topology optimization. Structural and Multidisciplinary Optimization, 41(5), 671-683.

[53] Martin, R. B., Burr, D. B., Sharkey, N. A., & Fyhrie, D. P. (1998). Skeletal tissue mechanics (Vol. 190). New York: springer.

[54] Hughes, T. J., Cottrell, J. A., & Bazilevs, Y. (2005). Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. Computer methods in applied mechanics and engineering, 194(39-41), 4135-4195.

[55] Delmarcelle, T., & Hesselink, L. (1994, October). The topology of symmetric, second-order tensor fields. In Proceedings Visualization'94 (pp. 140-147). IEEE.

[56] Hotz, I., Sreevalsan-Nair, J., Hagen, H., & Hamann, B. (2010). Tensor field reconstruction based on eigenvector and eigenvalue interpolation. In Dagstuhl Follow-Ups (Vol. 1). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

4 Application of redundancy paradigm to mechanisms and robotic manipulators

In the previous chapter, it had been presented a design tool which is based on the implementation of the directionality paradigm in order to modify a well-established topology optimization procedure. On the other hand, in this chapter, it is described the theoretical background and the actual application of the redundancy paradigm, inspired by biomimetic studies: the final goal is the proposal of an integrated design approach, or, in other words, the simultaneous definition of the ideal structure and control of a mechanism, in order to improve its performances.

Even if the outcomes of the two chapters are different, they share the same methodological background presented in the first one, and which is based on the two main pillars, getting inspiration from the observation of living beings, and translating the knowledge of biology principles in a mathematical form, and more precisely, in a constrained optimization problem.

What will be presented in the next sections is the base for the development of the architecture and control of innovative mechanism, inspired by the redundancy of the limbs of the animals, and which has been disclosed in an international patent application. Since a prototype of the mechanism does not exist yet, the validity of the results has been checked by the use of a virtual prototyping procedure, in order to simulate the behaviour of the mechanism and obtain a first ideal benchmark.

4.1 Use of the Digital Twin for the innovation of mechanisms and robotic manipulators

The goal of the virtual prototyping is the verification of the design of a product before the creation of an actual prototype. Usually, it is done using informatics tools able to provide simulations based on numerical approximations of the laws governing a physical system. There are different kinds of application which can be used as support of the design:

- it is possible to generate a geometrical model of the product, commonly used for dimensional and assembly control;

- it is possible to create a cinematic model of the physical system, typically for the design of mechanisms, useful for a first verification of functional requirements;
- static and dynamic are further and more complex models, and they take in account the forces acting on the system;
- it is possible to create a structural model of the product, determining the displacement field associated to the deformation of the continuum body, and the stress and strain tensor fields as well;
- implementing not only mechanical models, but even fluidic and electromagnetic models, it is possible to carry out a multi-physic analysis.

The higher is the number of phenomena described in the global model, the higher will be the complexity and the quality of the representation of the physical reality.

The setting of the model the simulation is based on, its informatics implementation, input and generated output data are the digital twin of the product or physical system. Starting from this definition, it is clear that the adoption of a feasible model is a crucial factor. A model is a conceptual representation of a phenomena, and, especially in engineering practice, it is simplified. In other words, it provides an approximation of the real system, neglecting minor aspects.

The scope of this section is illustrating the use of virtual models for the innovation of mechanisms and serial manipulators. In this case, the model must be able to describe the architecture, behaviour and control of the systems, based on the implementation of an innovative idea. More specifically, here an innovative solution is a specific design, or a control strategy generated by an inedited optimization problem. Such optimization process is defined starting by a new formulation of the design problem, removing some design specification for example, or removing some project boundaries.

Anyway, in order to verify the impact of a new design on the performances, or to solve the optimization problem itself, the virtual counterpart of the physical system must implement modified models, able to describe the new set up of the problem. As an example, which will be analysed in depth in all the present chapter, it is possible to take in consideration the inverse kinematic problem of a non-redundant serial manipulator. For this class of mechanisms, the inverse kinematic is a complicated problem, but the solution is a set of finite configurations of the system. But if a link is added to the mechanisms, it become a redundant manipulator, with infinite configurations compatible

with the prescribed pose of the terminal link. In general, to solve such problem, it is necessary to introduce a secondary task, in order to set up a determinate problem again. But if adding a degree of freedom takes to a mathematical indeterminacy, on the other hand it allows the mechanism to fulfil a new functional requirement.

4.2 Theoretical framework for serial mechanisms

4.2.1 Characterization of a serial manipulator and the Jacobian matrix

The Lagrange variables are the sufficient parameters describing the configuration of a mechanical system, such as a robotic manipulator, a 5-axis milling machine, or a 3D printer; if the system can be considered ideal, constituted by perfectly rigid bodies connected by kinematic couplings, these parameters correspond to the position of the actuators, both linear and angular. There is a force or a torque, provided by an actuator, in correspondence of every Lagrange variable.

There are many different mechanism typologies, and a first distinction is between a serial mechanism and a parallel mechanism. The former is characterized by a unique (vector) kinematic equation; on the contrary, the latter is characterized by a kinematic equation for every closed chain, plus the equation describing the pose of the terminal link.

In both cases, the kinematic analysis is based on the closure equations (both vector or scalar), of the kinematic chain of the mechanism. For instance, such equations are written in order to express the position of the end effector of the manipulator in function of the Lagrange variables. $\left(P_x, P_y, P_z\right)^T$ is the vector representing the position of the terminal link in the Euclidean space, and $\left(\Omega_{z3}, \Omega_{y4}, \Omega_{z5}\right)^T$ is the vector representing the pose, according to the Euler angles convention. Furthermore, $\underline{q} = (q_1, \dots, q_6)^T$ is the vector of the Lagrange variables describing the behaviour of the system. The closure equations are:

$$\begin{cases} P_x = P_x\left(\underline{q}\right) \\ P_y = P_y\left(\underline{q}\right) \\ P_z = P_z\left(\underline{q}\right) \end{cases} \leftarrow \begin{matrix} position\ of \\ the\ end-effector \end{matrix} \quad \begin{cases} \Omega_{z3} = \Omega_{z3}\left(\underline{q}\right) \\ \Omega_{y4} = \Omega_{y4}\left(\underline{q}\right) \\ \Omega_{z5} = \Omega_{z5}\left(\underline{q}\right) \end{cases} \leftarrow \begin{matrix} angular\ pose\ of \\ the\ end-effector \end{matrix} \tag{4.1}$$

in a more compact way (vector notation):

$$\underline{P} = \left(P_x, P_y, P_z, \Omega_{z3}, \Omega_{y4}, \Omega_{z5}\right)^T = \underline{P}(q_1, \dots, q_6) = \underline{P}\left(\underline{q}\right) \tag{4.2}$$

Starting from the closure equations, one key tool for the description of the behaviour of the system is the Jacobian matrix:

$$\begin{vmatrix} \dot{P}_x \\ \dot{P}_x \\ \dots \\ \dot{\Omega}_{z5} \end{vmatrix} = \begin{vmatrix} \\ \underline{\underline{J}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \dot{q} = \begin{vmatrix} \\ \underline{\underline{J}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \begin{vmatrix} \dot{q_1} \\ \dot{q_2} \\ \dots \\ \dot{q_6} \end{vmatrix} \tag{4.3}$$

$\underline{\underline{J}}(\underline{q})$ describes the variation of the position and angular pose of the end effector with respect to the Lagrange variables [57]. In this case the mechanism owns 6 actuators, and its primary task is imposing a certain position and a certain orientation for the final link. In other words, characterizing the position of the end effector in function of the angular positions of its joints is equivalent to defining a mapping $\mathbb{R}^6 \rightarrow \mathbb{R}^6$ between the Lagrange variables space, or joints space, and the Cartesian space.

In general, there are two Jacobian matrixes: the geometric Jacobian, and the analytic Jacobian. The former defines a relation between the linear and angular velocity of the end effector referred to the time variation of the Lagrange variables. The latter is the one which has been introduced above, and it refers to the minimal form description of the mechanism: differentiating both sides of such relation it is possible to obtain the analytical Jacobian. In the case of a plane problem, the two matrixes are equivalent, and this shall be an important remark to take in consideration when it will be introduced the case study of the redundant workpiece turntable.

4.2.2 Characterization of a serial manipulator trajectory

In general, the primary task of a manipulator is positioning the end effector on different points, belonging a given line, called path, with a certain time law: the set composed by the path and the time law is called trajectory.

The path is the locus of the consecutive positions of the end-effector, and it is a one-dimension variety in the cartesian space $\mathbb{R}^3$, and its points $\underline{R} = \left(R_x, R_y, R_z\right)^T$ can be writen in function of a unique parameter:

$$\begin{cases} R_x = R_x(s) \\ R_y = R_y(s) \\ R_z = R_z(s) \end{cases} \tag{4.4}$$

In the special case of a linear path, these are proportionality relations with 3 constants:

$$\begin{cases} R_x = s \cdot k_1 = s \cdot cos\alpha \\ R_y = s \cdot k_2 = s \cdot cos\beta \\ R_z = s \cdot k_3 = s \cdot cos\gamma \end{cases} \tag{4.5}$$

In the equations (4.5) the constants are evidently the three direction cosines with $cos^2\alpha + cos^2\beta + cos^2\gamma = 1$, and the s parameter represents the curvilinear coordinate, or, in other words, the length of the arc when traveling from the origin to the given point $\underline{R}(s)$. The same reasoning is valid for the pose of the end effector, even if it is slightly more difficult to visualize:

$$\begin{cases} \Omega_{z3} = \Omega_{z3}(s) \\ \Omega_{y4} = \Omega_{y4}(s) \\ \Omega_{z5} = \Omega_{z5}(s) \end{cases} \tag{4.6}$$

On the other hand, taking in account the time law, it is necessary to describe the dependency of the position of the end effector in function of the time, or, formally, determine the following:

$$\begin{cases} R_x = \hat{R}_x(t) \\ R_y = \hat{R}_y(t) \\ R_z = \hat{R}_z(t) \\ \Omega_{z3} = \hat{\Omega}_{z3}(t) \\ \Omega_{y4} = \hat{\Omega}_{y4}(t) \\ \Omega_{z5} = \hat{\Omega}_{z5}(t) \end{cases} \tag{4.7}$$

One way to express the system (4.4) is using the curvilinear coordinate: if a function $s = s(t)$ is defined, then it can be written:

$$\begin{cases} \hat{R}_x(t) = R_x(s(t)) = R_x(s) \\ \hat{R}_y(t) = R_y(s(t)) = R_y(s) \\ \hat{R}_z(t) = R_z(s(t)) = R_z(s) \\ \hat{\Omega}_{z3}(t) = \Omega_{z3}(s(t)) = \Omega_{z3}(s) \\ \hat{\Omega}_{y4}(t) = \Omega_{y4}(s(t)) = \Omega_{y4}(s) \\ \hat{\Omega}_{z5}(t) = \Omega_{z5}(s(t)) = \Omega_{z5}(s) \end{cases} \tag{4.8}$$

Deriving by time, and using the chain derivation rule:

$$\begin{cases} \frac{d\hat{R}_x}{dt} = \frac{dR_x}{ds} \cdot \frac{ds}{dt} = \frac{dR_x}{ds} \cdot \dot{s} \\ \frac{d\hat{R}_y}{dt} = \frac{dR_y}{ds} \cdot \frac{ds}{dt} = \frac{dR_y}{ds} \cdot \dot{s} \\ \frac{d\hat{R}_z}{dt} = \frac{dR_z}{ds} \cdot \frac{ds}{dt} = \frac{dR_z}{ds} \cdot \dot{s} \\ \frac{d\hat{\Omega}_{z3}}{dt} = \frac{d\Omega_{z3}}{ds} \cdot \frac{ds}{dt} = \frac{d\Omega_{z3}}{ds} \cdot \dot{s} \\ \frac{d\hat{\Omega}_{z3}}{dt} = \frac{d\Omega_{y4}}{ds} \cdot \frac{ds}{dt} = \frac{d\Omega_{y4}}{ds} \cdot \dot{s} \\ \frac{d\hat{\Omega}_{z5}}{dt} = \frac{d\Omega_{z5}}{ds} \cdot \frac{ds}{dt} = \frac{d\Omega_{z5}}{ds} \cdot \dot{s} \end{cases} \tag{4.9}$$

Let us now consider the first three equations of the system (4.9): the vector $\left(\frac{dR_u}{ds}, \frac{dR_y}{ds}, \frac{dR_w}{ds}\right)^T$ is tangent to the one-dimension variety in each point. Furthermore, keeping in mind the definition of velocity of a point moving along a path:

$$\underline{V} = \frac{d\underline{\hat{R}}}{dt} = \begin{vmatrix} \frac{d\hat{R}_x}{dt} \\ \frac{d\hat{R}_y}{dt} \\ \frac{d\hat{R}_z}{dt} \\ \frac{d\hat{\Omega}_{z3}}{dt} \\ \frac{d\hat{\Omega}_{y4}}{dt} \\ \frac{d\hat{\Omega}_{z5}}{dt} \end{vmatrix} = \begin{vmatrix} \frac{dR_x}{ds} \cdot \dot{s} \\ \frac{dR_y}{ds} \cdot \dot{s} \\ \frac{dR_z}{ds} \cdot \dot{s} \\ \frac{d\Omega_{z3}}{ds} \cdot \dot{s} \\ \frac{d\Omega_{y4}}{ds} \cdot \dot{s} \\ \frac{d\Omega_{z5}}{ds} \cdot \dot{s} \end{vmatrix} = \begin{vmatrix} \frac{dR_x}{ds} \\ \frac{dR_y}{ds} \\ \frac{dR_z}{ds} \\ \frac{d\Omega_{z3}}{ds} \\ \frac{d\Omega_{y4}}{ds} \\ \frac{d\Omega_{z5}}{ds} \end{vmatrix} \cdot \dot{s} = \underline{\mathcal{L}} \cdot \dot{s} \tag{4.10}$$

If the path $\underline{R}(s) = (R_u(s), R_v(s), R_w(s))^T$ is defined in the way that s is actually the curvilinear coordinate, then the vector expression of a point along the path is the product

of the scalar velocity of the point in the curvilinear system $\dot{s}$, times the verso of the vector $\left(\frac{dR_u}{ds}, \frac{dR_y}{ds}, \frac{dR_w}{ds}\right)^T$.

4.2.3 Jacobian matrix and cinematic static duality

Once the kinematic behaviour equations of a serial manipulator are defined, it is possible to derive the static behaviour of the system straight forward. The goal is determining the forces and the torques required to the actuator and motors, when the system is in equilibrium with the forces and torques applied to the end effector. This is possible because the so called kinematic-static duality exists for the mechanisms composed by rigid bodies connected by ideal kinematic couples, and it is a direct consequence of the principle of the virtual works.

As previously reported, it exists the relation (4.3)

$$\begin{vmatrix} \dot{P}_x \\ \dot{P}_y \\ \dot{P}_z \\ \omega_x \\ \omega_y \\ \omega_z \end{vmatrix} = \begin{vmatrix} \\ \underline{\underline{\hat{J}}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \dot{q} = \begin{vmatrix} \\ \underline{\underline{J}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \begin{vmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dots \\ \dot{q}_6 \end{vmatrix}$$

which can be written in the more generic form

$$\begin{vmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \dots \\ \dot{p}_m \end{vmatrix} = \begin{vmatrix} \\ \underline{\underline{\hat{J}}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \dot{q} = \begin{vmatrix} \\ \underline{\underline{J}}(\underline{q}) \\ \\ \end{vmatrix} \cdot \begin{vmatrix} \dot{q}_1 \\ \dot{q}_1 \\ \dots \\ \dot{q}_n \end{vmatrix} \qquad (4.11)$$

with $\underline{q} = (q_1, \dots, q_6)^T$ being the vector of the joints variables (lagrangian variables) necessary for the description of the configuration of the system. The q_i parametrs can represent both rotations or traslations, depending on the kind of actuators or couples. If the mechanism is redundant, the degree of redundancy is $(n - m)$.

Differently by the previous subsection, the symbol $\underline{\underline{\hat{J}}}(\underline{q})$ represent the geometric Jacobian matrix, which is actually the relation between the time derivative of the Lagrange variables, and the set of both linear velocity $\left(P_x, P_y, P_z\right)^T$ of the origin of the end-effector reference system, and the angular velocity of the end effector reference system with respect to the ground frame.

In the case of the 6 degrees of freedom serial manipulator, such as an anthropomorphic manipulator with a spherical wrist, it is possible to define the vector $\left(F_x, F_y, F_z\right)^T$ of the forces, and the vector $\left(M_x, M_y, M_z\right)^T$ of the torques applied to the wrist, with to respect the ground reference system. Furthermore, the vector of the forces or torques provided by the actuator are stored in the vector $(\gamma_1, \gamma_2, \dots, \gamma_6)^T$, having the same indexes order as the vector of the Lagrange variables. Under this hypothesis:

$$\begin{vmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_6 \end{vmatrix} = \begin{vmatrix} \\ \underline{\underline{\hat{J}}}^T_{(6x6)}(q_1, q_2, .., q_6) \\ \\ \end{vmatrix} \cdot \begin{vmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{vmatrix} \tag{4.12}$$

More generally, defining a vector $(f_1, f_2, \dots, f_m)^T$ of m generalized forces, and a set $(q_1, \dots, q_n)^T$ of n lagriangian variables, then it can be written:

$$\begin{vmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_n \end{vmatrix} = \begin{vmatrix} \\ \underline{\underline{\hat{J}}}^T(q_1, q_2, .., q_n) \\ \\ \end{vmatrix} \cdot \begin{vmatrix} f_1 \\ f_2 \\ \\ f_{\mathrm{m}} \end{vmatrix} \tag{4.13}$$

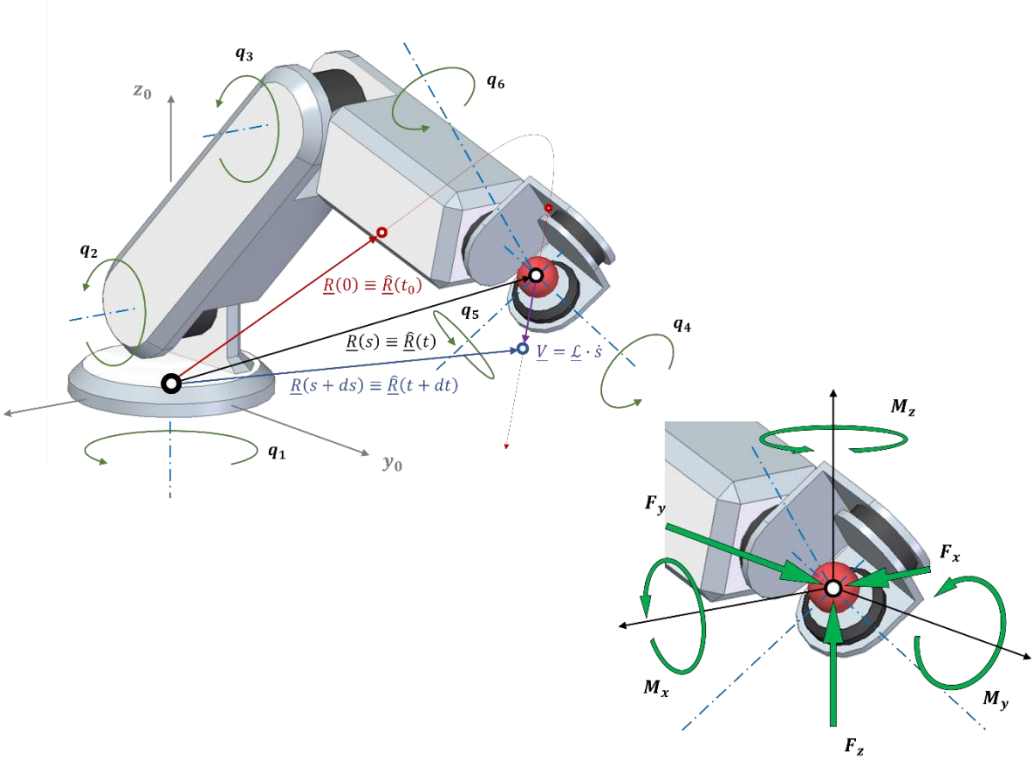


Figure 4.1: representation of the anthropomorphic manipulator with a spherical wrist

As an example, let us consider a certain position of the end effector, and a set of forces acting on it: if the mechanism is able to minimize the sum of the torques required to the actuators, it means that the mechanism is redundant, because the quantity $(n - m) > 0$, and the equations providing the actuator torques $(\gamma_1, \gamma_2, \dots, \gamma_n)^T$ are candidates for the creation of an objective function of the optimization problem.

4.2.4 Arbitrariness of the primary task and definition of the actual Jacobian matrix

Starting from the considerations done previously, it is possible to figure out that it is necessary introduce a degree of redundancy in the mechanism in order to allow the definition of a secondary task for the mechanism, beside the primary positioning requirement. Anyway, somehow, the previous sentence is not completely true, because a mechanism is not intrinsically redundant or not, but such qualification depends on the description of the tasks it must fulfil, and the control parameters. Let's consider an anthropomorphic manipulator with a spherical wrist: in the case of complete definition of

position and pose of the end effector, 6 parameters are required, and the Jacobian matrix can be written as:

$$\begin{cases} \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{d\Omega_{z3}\left(\underline{q}\right)}{dt} = \frac{d\Omega_{z3}(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial \Omega_{z3}}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial \Omega_{z3}}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial \Omega_{z3}}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{d\Omega_{y4}\left(\underline{q}\right)}{dt} = \frac{d\Omega_{y4}(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial \Omega_{y4}}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial \Omega_{y4}}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial \Omega_{y4}}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{d\Omega_{z5}\left(\underline{q}\right)}{dt} = \frac{d\Omega_{z5}(\,q_1,\dots,q_6\,)}{dt} = \frac{\partial \Omega_{z5}}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial \Omega_{z5}}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial \Omega_{z5}}{\partial q_6}\cdot\frac{dx_6}{dt} \end{cases} \tag{4.14}$$

or, in a matrix form:

$$\underline{\dot{P}} = \begin{vmatrix} \dot{P}_x \\ \dot{P}_y \\ \dot{P}_z \\ \dot{\Omega}_{z3} \\ \dot{\Omega}_{y4} \\ \dot{\Omega}_{z5} \end{vmatrix} = \begin{vmatrix} \frac{dP_x}{dt} \\ \frac{dP_y}{dt} \\ \frac{dP_z}{dt} \\ \frac{d\Omega_{z3}}{dt} \\ \frac{d\Omega_{y4}}{dt} \\ \frac{d\Omega_{z5}}{dt} \end{vmatrix} = \begin{vmatrix} \frac{\partial P_x}{\partial q_1} & \frac{\partial P_x}{\partial q_2} & \cdots & \frac{\partial P_x}{\partial q_6} \\ \frac{\partial P_y}{\partial q_1} & \frac{\partial P_y}{\partial q_2} & \cdots & \frac{\partial P_y}{\partial q_6} \\ \frac{\partial P_z}{\partial q_1} & \frac{\partial P_z}{\partial q_2} & \cdots & \frac{\partial P_z}{\partial q_6} \\ \frac{\partial \Omega_{z3}}{\partial q_1} & \frac{\partial \Omega_{z3}}{\partial q_2} & \cdots & \frac{\partial \Omega_{z3}}{\partial q_6} \\ \frac{\Omega_{y4}}{\partial q_1} & \frac{\Omega_{y4}}{\partial q_1} & \cdots & \frac{\Omega_{y4}}{\partial q_1} \\ \frac{\Omega_{z5}}{\partial q_1} & \frac{\Omega_{z5}}{\partial q_1} & \cdots & \frac{\Omega_{z5}}{\partial q_1} \end{vmatrix} \cdot \begin{vmatrix} \dot{q}_1 \\ q_2 \\ \dot{q}_3 \\ \dot{q}_3 \\ \dot{q}_3 \\ \dot{q}_3 \end{vmatrix} = \underline{\underline{J}}_{(6x6)} \cdot \underline{\dot{q}} \tag{4.15}$$

This formulation is convenient if we use the mechanism as a manipulator. But let us suppose that a tool is positioned on the terminal link, in order to use the mechanism as a milling machine, and let us suppose furthermore that the tool axis is coincident with one of the axes of the wrist, let's say the z_5 axis. In this case, the number of the degrees of freedom which must be defined is five, because the rotation of the terminal link around the z_5 axis has no influence on the primary task, which is the correct positioning of the tool with respect the workpiece. In this particular situation, the closure equations for the kinematic chain (4.1) become:

$$\begin{cases} P_x = P_x\left(\underline{q}\right) \\ P_y = P_y\left(\underline{q}\right) \\ P_z = P_z\left(\underline{q}\right) \end{cases} \leftarrow \begin{matrix} position \\ of\ the\ tool \end{matrix} \quad \begin{cases} \Omega_{z3} = \Omega_{z3}\left(\underline{q}\right) \\ \Omega_{y4} = \Omega_{y4}\left(\underline{q}\right) \end{cases} \leftarrow \begin{matrix} angular\ pose \\ of\ the\ tool \end{matrix} \tag{4.16}$$

Consequently, the Jacobian matrix becomes:

$$\begin{cases} \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(q_1,\dots,q_6)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(q_1,\dots,q_6)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{dP_x\left(\underline{q}\right)}{dt} = \frac{dP_x(q_1,\dots,q_6)}{dt} = \frac{\partial P_x}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial P_x}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial P_x}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{d\Omega_{z3}\left(\underline{q}\right)}{dt} = \frac{d\Omega_{z3}(q_1,\dots,q_6)}{dt} = \frac{\partial \Omega_{z3}}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial \Omega_{z3}}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial \Omega_{z3}}{\partial q_6}\cdot\frac{dx_6}{dt} \\ \frac{d\Omega_{y4}\left(\underline{q}\right)}{dt} = \frac{d\Omega_{y4}(q_1,\dots,q_6)}{dt} = \frac{\partial \Omega_{y4}}{\partial q_1}\cdot\frac{dq_1}{dt} + \frac{\partial \Omega_{y4}}{\partial q_2}\cdot\frac{dx_2}{dt} + \dots + \frac{\partial \Omega_{y4}}{\partial q_6}\cdot\frac{dx_6}{dt} \end{cases} \tag{4.17}$$

or, in matrix form:

$$\underline{\dot{P}} = \begin{vmatrix} \dot{P}_x \\ \dot{P}_y \\ \dot{P}_z \\ \dot{\Omega}_{z3} \\ \dot{\Omega}_{y4} \end{vmatrix} = \begin{vmatrix} \frac{dP_x}{dt} \\ \frac{dP_y}{dt} \\ \frac{dP_z}{dt} \\ \frac{d\Omega_{z3}}{dt} \\ \frac{d\Omega_{y4}}{dt} \end{vmatrix} = \begin{vmatrix} \frac{\partial P_x}{\partial q_1} & \frac{\partial P_x}{\partial q_2} & \dots & \frac{\partial P_x}{\partial q_6} \\ \frac{\partial P_y}{\partial q_1} & \frac{\partial P_y}{\partial q_2} & \dots & \frac{\partial P_y}{\partial q_6} \\ \frac{\partial P_z}{\partial q_1} & \frac{\partial P_z}{\partial q_2} & \dots & \frac{\partial P_z}{\partial q_6} \\ \frac{\partial \Omega_{z3}}{\partial q_1} & \frac{\partial \Omega_{z3}}{\partial q_2} & \dots & \frac{\partial \Omega_{z3}}{\partial q_6} \\ \frac{\Omega_{y4}}{\partial q_1} & \frac{\Omega_{y4}}{\partial q_1} & \dots & \frac{\Omega_{y4}}{\partial q_1} \end{vmatrix} \cdot \begin{vmatrix} \dot{q}_1 \\ q_2 \\ \dot{q}_3 \\ \dot{q}_3 \\ \dot{q}_3 \\ \dot{q}_3 \end{vmatrix} = \underline{\underline{J}}_{(5x6)} \cdot \underline{\dot{q}} \tag{4.18}$$

The Jacobian matrix competing the anthropomorphic manipulator with the spherical wrist for the task of robotic machining is a 5x6 matrix, which implies that the mechanism has a redundant degree of freedom. This means that, assigning a new primary task to the mechanism allows the introduction of a secondary task, or, more precisely, it is necessary define a secondary task, in order to solve the control problem: as an example, it is possible to impose that the configuration of the mechanism is such that it minimizes the torques at the joint actuators, in order to improve the stiffness of the whole mechanism. The definition of the ideal configuration of the manipulator requires the solution of an

optimization problem. Typically, the solution algorithm of the optimization problem is implemented in the digital twin of the physical system.

4.3 Definition of the optimization problem for the innovation of the mechanisms

In the previous section it had been highlighted that it is possible to make the positioning task indeterminate by introducing a redundant degree of freedom in the mechanism architecture or by modifying the primary positioning task. Once the problem associated to the pose of the manipulator is indeterminate, it is necessary to introduce a secondary task which is associated to an optimization problem.

As an example, let us consider a 5-axis milling machine. It is well known by literature that positioning errors of the tool with respect to the work piece can be partially explained by the Abbe principle [58]. Actually, the volumetric errors are caused by a propagation of the dimensional and assembly errors of the machine, along its kinematic chain. As it will be better described in the next section, it is possible to partially overcome the Abbe effect by introducing a redundant actuator in the architecture of the machine, and control the milling process minimizing the distance between the position of the tool, and the rotation centre of the work piece table. In other words, it is possible to set a minimization problem such as:

$$\begin{cases} (\mathbf{1}) \quad \boldsymbol{min}.: \quad influence\ of\ dimensonal\ errors\ on\ volumetric\ errors \\ (\mathbf{2}) \quad \boldsymbol{s.t.}: \quad closure\ equations\ (posizionamneto\ end\ effector) \\ (\mathbf{3}) \quad \boldsymbol{s.t.}: \quad time\ law\ of\ the\ trajectory \end{cases} \tag{4.19}$$

Similarly to chapter 3 regarding TO, here there is an objective function which must be minimized (4.19(1)), and some constraints (4.19(2) and 4.19(3)). The latter are constitutive equations which describe the physical model, or equations describing the control logic of the mechanism: these are the equations the virtual model is based on.

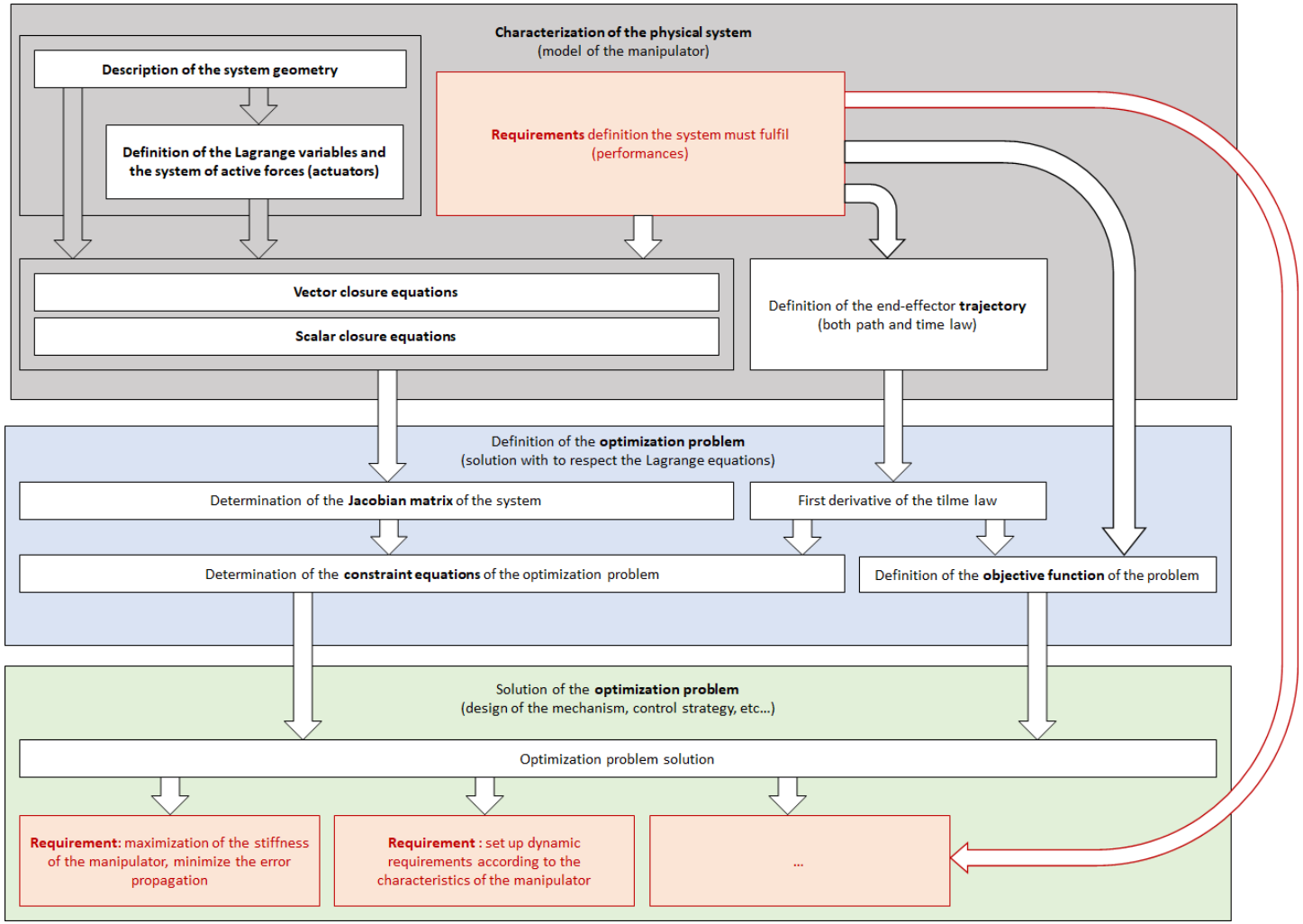


Figure 4.2: definition of the optimization problem based on the model of the physical system

In the system (4.19), the lower is the value of the objective function, the higher are the performances of the mechanism. Other objective functions which may be defined are:

- minimization of the required couples and forces provided by motors and actuators;
- restraint of the dynamic performances required to the manipulator;
- expansion of the mechanism workspace.

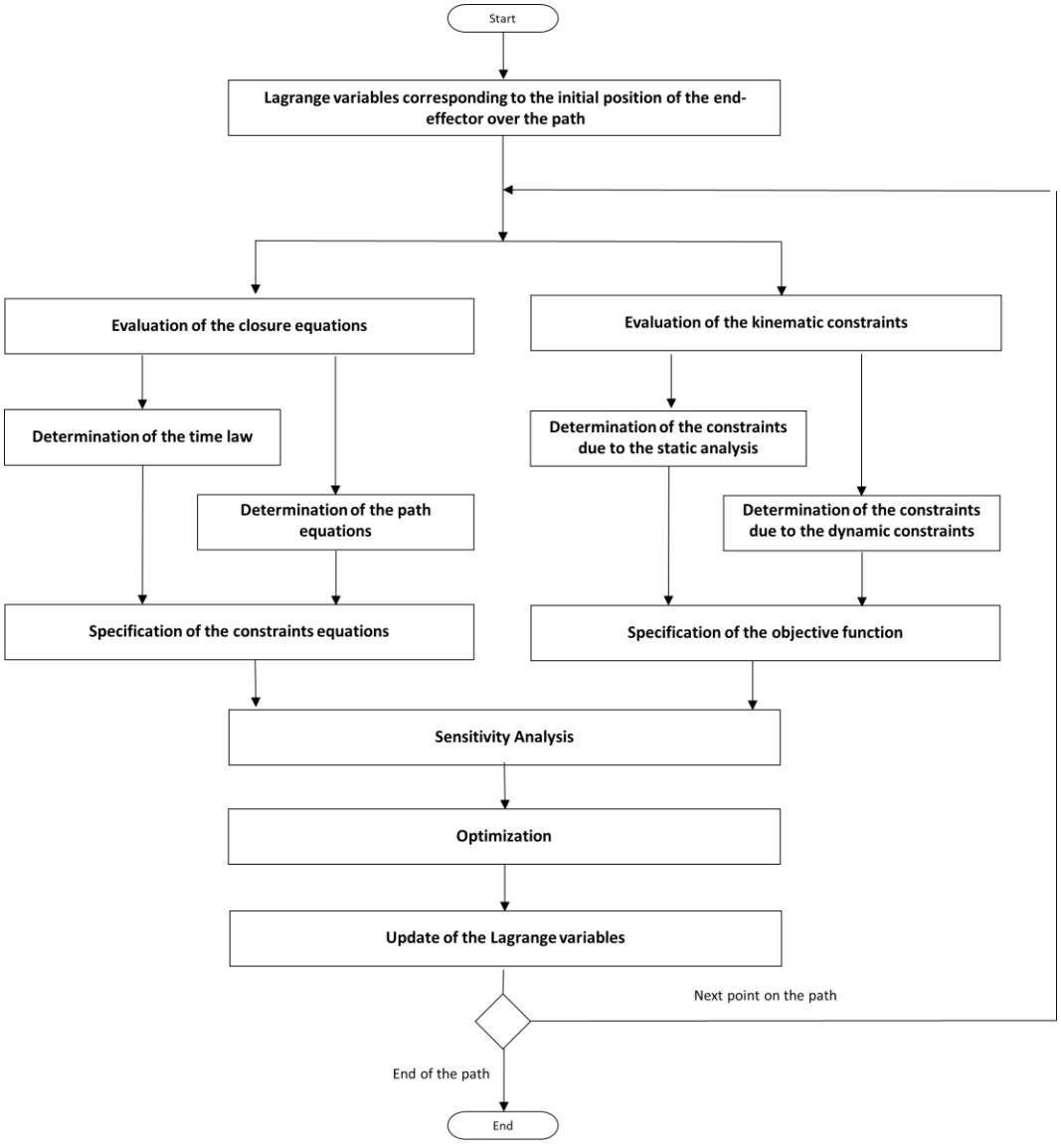


Figure 4.3: diagram of the pseudo-algorithm of the optimization process of the performance a redundant serial manipulator

A schema depicting the main features which must be implemented in the serial mechanism digital twin for the performance optimization is depicted in figure 4.2.

Furthermore, in the figure 4.3 is depicted the block diagram for the secondary task (stiffness improvement, positioning errors minimization, etc…), taking in account the boundaries provided by the primary task (keep the position of the end effector on the trajectory, according with the motion laws).

4.4 An actual application of the Redundancy Paradigm

The schema described in the previous sections offers a methodological framework for the concept of a new device for the milling and 3D printing. The main idea has been exploited in one national and one international patent application [59], and a technical insight has been provided in a journal article which describes the theoretical aspects [60]. In this section it will be firstly presented the relevant state of the art, and then a brief resume of the general idea will be provided. Moreover, there will be introduced the theoretical background and the study of the proposed redundant mechanism by the means of some simulations. Finally, all the improvements will be showed and discussed, in order to evaluate the effectiveness of the proposed methodology.

4.4.1 State of the art

Machine tools, such as 5 axis milling machines, are devices able to produce mechanical components by the mean of the elimination of portions of material from a starting workpiece. The effective material removal is caried out by the tool, that must assume a prescribed position, and a prescribed angular pose, with the higher possible precision. A low accuracy in fulfilling this task leads to obtain mechanical pieces affected by dimensional and shape errors. An error is defined as the norm of the vector representing the distance between the ideal configuration of the machine, and the actual position of the tool respect to the finished part reference frame. The deviation of the tool placement from the prescribed one depends by different factors, but it always directly causes the unconformity of the dimensions of the work piece with to respect nominal specifics.

Many authors [61] refer to this quantity as volumetric error, which represents an index for the errors of a machine tool [62].

Geometric unconformity of machine's components, assembly errors, thermal conditions, wear, are all sources of volumetric errors, and many researchers proposed different approaches for the modelling of the different aspects. In order to study the propagation of the dimension unconformity of the components along the kinematic chain of the mechanism, it had been proposed the use of Denavit and Hartenberg convention [63][64], or the one proposed by Suh and Lee and Jung [65]. Another, and more recent example has been provided by Cheng et al. [66]. Anyway, in general, it is possible to state that the first efforts in providing reliable methods for the prediction of volumetric errors focused mainly on the study of three axis machines [67], due to the difficult in considering long kinematic chains comprising rotational couples. Later on, models for the identification of errors cause by angular deviation have been developed [68], and it is remarkable how the topic is still of great interest, and still an open field [69].

The mitigation of the geometric error propagation along the kinematic chain of the mechanism is the final goal, despite the fact that many different approaches can be applied for its determination. It can be said that, in the design of machine tools, the adopted strategies are mainly of two kinds: the first one is the error compensation [70], which implies active countermeasures based on real time measurements, integrated in the control logic. On the other hand, error avoidance [71] has been taken in account, which prescribes the mitigation of the influence of geometric uncertainties by the intrinsic use of particular design principles: these may be the adoption of particular constructive material, or the adoption of a particular morphology for the structure.

Error mitigation is the goal of the implementation of the method presented in this chapter, since the proposed modify of control logic is not based on contextual measurements, but only on the management of the redundancy, which intrinsically allows the minimization of the position errors, or the optimization of other performance indexes.

As it had been previously explained, redundancy is not an intrinsic feature of a mechanism, but it is related to its functional requirements. For a 5 axis machines tools, the objective is positioning the tool in a certain position and with a certain slope respect to the work piece. This means that the 5 degrees of freedom of the tool (the rotation

around the symmetry axis is not taken in account) must be defined with respect to the reference frame attached to the work piece.

For this reason, usually the milling machines are characterized by a kinematic chain composed by five links connected by five actuators or motors. As it will be introduced later, the RRTTT axis milling machine, as defined by Kiridena and Ferreira [72], is a typical architecture based on this schema. An alternative redundant architecture is here proposed, according to the method described by the previous subsection.

Other examples of redundant mechanism for manufacturing purposes may be found in in literature, and the use of industrial robots is an alternative to traditional CNC machines [73][74], has been already disclosed.

For instance, a standard industrial robot may be characterized by 3 actuators belonging to the anthropomorphic manipulator, and 3 belonging to the spherical wrist. With 6 actuated degrees of freedom industrial robots have a redundant degree of freedom compared to the 5 DOF required for a classical milling operation. This provides a higher flexibility of the manipulator, which can be used in order to improve the rigidity of the system during the machining [75].

Actually, the most interesting features of the redundant systems is the possibility to deal with different tasks at the same time [59]. A possible secondary task, as disclosed by Xiao and Huan [76] is the avoidance of singularities, the respect of the joint limits, and the prevention of the collisions. Furthermore, another example is the maximization of the resulting global stiffness of the manipulator depending on the configuration.

The analysis that will be presented below focuses on a variant of the RRTTT 5 axis machine tool: this production means may be described as mechanisms equipped with a tilting table, having two rotational degrees of freedom (RR); the table has the task of supporting the workpiece; on the other hand, the tool is characterized by tree translational degrees of freedom (TTT). Starting from this configuration, the redundant TRRTTT milling machine is derived, while a further translation degree of freedom is mounted upon the tilting table, as it is depicted in figure 4.6. Taking advantage of the redundant configuration, it is possible to determine the ideal configuration of the mechanism according to the achievement of a secondary task.

The volumetric error, caused by the deviation of all the dimensional parameters from their ideal values, is the first objective function taken in account. Ideally, it should be necessary

to take in account the dimensional errors of all the components of the machine, as well as the assembling unconformities. In practice, the focus here is on the angular errors introduced by the table motors because, as it had been highlighted by Chen et al. [77], roll pitch and yaw errors have a high influence compared to linear positioning. This can be explained by the Abbe effect, which postulate an amplification of angular uncertainties depending on the structure of the kinematic chain. The introduction of a redundant degree of freedom is adopted in order to mitigate the effect of the angular uncertainties introduced by the tilting table. More precisely, a translational axis is added to the double turntable, as it will be explained more in detail later. Furthermore, a second objective function will be taken in account, and more precisely the objective will be the minimization of the torque required to the motor of the turntable during the machining of the workpiece.

The use of a virtual model of the manufacturing machine is proposed here as a tool for evaluating the improvement of the performances due to the introduction of the new axis in the modified architecture. Two goals can be achieved simultaneously: the first one is an investigation purpose, and the second is the design of the control. Regarding the first point, as it had been reported by Pedersen et al. [78], a digital twin of a physical system is an effective tool for studding the behaviour of the actual mechanism. On the other hand, the definition of virtual model of the mechanism is a necessary step for the implementation of the optimization problem in the control. In fact, the fulfilment of the optimality conditions will allow to obtain the improvement of the machine performances.

A last remark is about the validity of the framework for other types of devices. In fact, additive manufacturing means would benefit of the minimization of the placement error of the nozzle, or the torch: their correct position with respect to the workpiece is a key factor for the correct generation of the product, in order to minimize the deviation from the original design [79][80][81]. Consequently, the importance of mitigating the influence of geometric uncertainties of the productive process is still valid in productive means adopting both subtractive and additive paradigms, such as hybrid machines [82].

4.4.2 Main idea

By the analysis of the state of the art, it is clear that one of the main issues, in the fabrication of mechanical components, is the possibility of having a degradation of the quality of the manufacturing process during the manufacture of a work piece. This is true in milling process and in additive manufacturing procedures as well. Such phenomena may have different causes, since the production of mechanical component is an articulated activity, and production means are devices with a high level of complexity. As it had been highlighted, one of the possible factors is the Abbe effect, which describes the introduction of errors in the machining of a work piece as an amplification of dimensional and assembly errors of the production mean along its kinematic chain.

In order to provide an example, let's consider the simplified schema of a production process depicted in figures 4.4, that, actually, can be a milling process, or an additive manufacturing process. The system has 3 degrees of freedom, two provided by linear actuators, and one provided by a rotary motor. The prismatic kinematic couples provide the positioning of the nozzle/tool with respect to the work piece, and the motor impose the slope of the work piece. During the machining of the work piece, the position of the nozzle/tool with to respect the centre of rotation of the rotary kinematic couple changes. An example of this fact is provided by the confront of figures 4.4(a) and 4.4(b). This relative position is unique because the mechanism is not redundant to the task of positioning the nozzle/tool with respect the work piece. Typically, the distance between the tool tip and the piece increases during the production process, and, according to the Abbe effect, this takes a deterioration of the quality of the production process. This phenomenon is shown in figures 4.4(c) and 4.4(d), where it can be noticed the increment of the vector $\underline{\Delta} = (\Delta x, \Delta y)$, which represents contribution to the volumetric error introduced by the angular error.

In order to evaluate the loss of quality, let us consider the relative position of the tip and the work piece as a function of the table rotation θ: the final goal is expressing the sensitivity of the relative position as a function of the distance. Let L be the distance between the tip and the centre of rotation, then a slope variation of the angle θ will produce a variation $Lsin\theta$ of the relative position. If we consider a small perturbation ε of the slope it is possible to write:

$$\frac{\partial x}{\partial \theta} = \frac{\partial}{\partial \theta}\left(r\sin(\theta+\varepsilon)\right) = r\cos(\theta+\varepsilon) \tag{4.20}$$

where x is the horizontal component of the relative displacement, and with the hypothesis that the error ε does not depend by θ. The higher influence of the angular positioning error is in correspondence of the value $\theta = 0$, according with the equation (4.20), and the error amount is:

$$\Delta x = r\sin(\theta+\varepsilon) - r\sin\theta = r(\sin(\theta+\varepsilon) - \sin\theta) \tag{4.21}$$

The sensitivity of the error Δx to the angular position error is linear with respect to the relative distance between the tip and the work piece. Looking at the equation (4.21), it is easy to notice that there are two ways to nullify the error: impose $\varepsilon = 0$, or impose $r = 0$. Obviously, ε can not be nullified because is the angular error, and L can not be controlled because it is consequence of the closure of the kinematic chain.

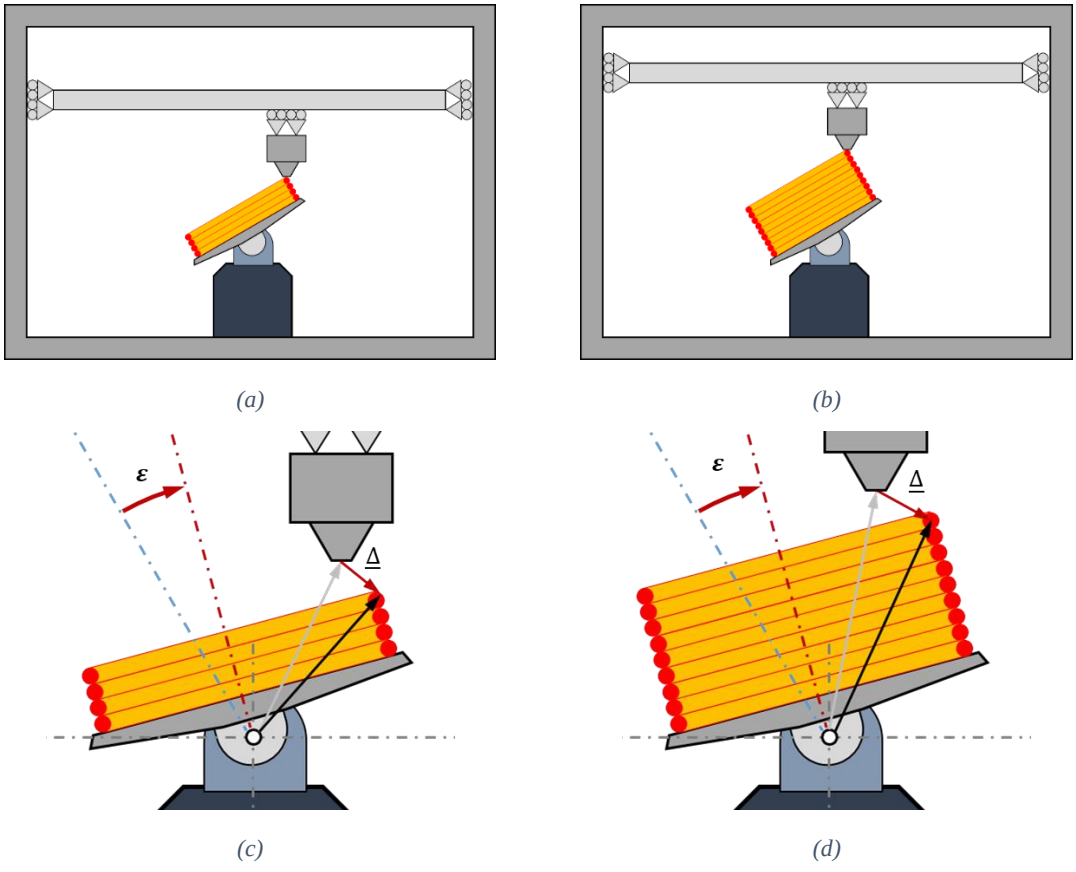


Figure 4.4: representation of two different phases of the machining process of a component using a standard additive manufacturing device

In order to improve the performance of the mechanism in terms of accuracy of relative position between the tool and the work piece, it is possible to adopt the methodology introduced in the previous sections. Actually, according to the schema in figures 4.5, it is possible to add a linear actuator to the architecture of the work piece platform. In this way, the 3D printer or the milling machine become redundant to the primary positional task, and it is possible to introduce a secondary task, which, in this case, is the minimization of the distance between the tip of the tool and the rotation centre of the table.

Under the said conditions, the formulation of the optimization problem, which is generalized according to the system (4.19), becomes:

$$\begin{cases} \boldsymbol{Objective\ function}\ (secondary\ task\ of\ the\ mechanism): \\ minimization\ of\ the\ distance\ L\ between\ the\ tool\ and\ the\ rotation\ center \\ \\ \boldsymbol{Boundary\ conditions}\ (primary\ task\ of\ the\ mechanism): \\ realize\ the\ correct\ position\ of\ the\ work\ piece\ with\ to\ respect\ the\ tool\ tip \\ according\ with\ the\ machining\ process \end{cases} \tag{4.22}$$

The implementation of the secondary task makes the positioning problem determinate again, but, differently by the standard machine, the redundant degree of freedom allows the reduction of the sensitivity of the position errors with to respect the machine defects (wrong angular position of the table motor).

Figures 4.5(a) and 4.5(b) depict two different control strategies: in figure 4.5(a), the redundant axis is not controlled, and the configuration of the machine is actually equivalent to the not redundant one, shown in figure 4.4(b); consequently, the error vector reported in figure 4.5(c) will have the same module as the one in figure 4.4(d). On the contrary, if the optimization strategy is implemented, the redundant prismatic axis is controlled in way that minimizes the distance between the rotation point of the table, and position of the nozzle or tool tip: as a consequence, the volumetric error is minimized, according to equation (4.21).

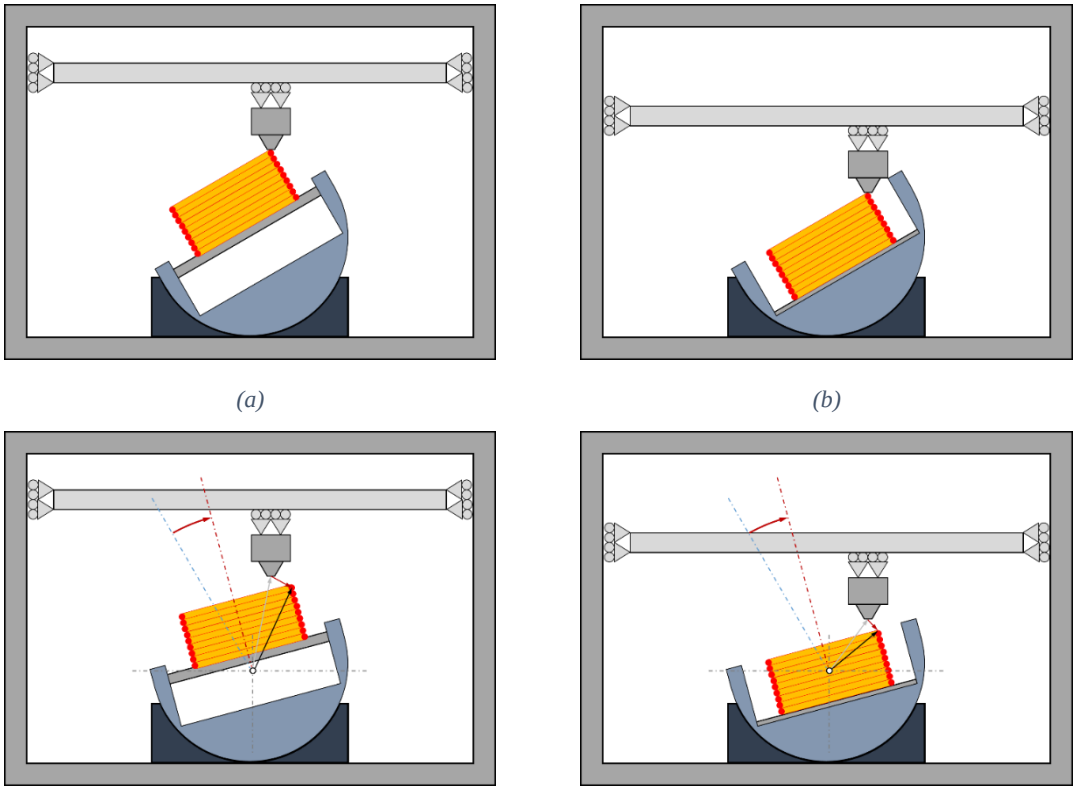


Figure 4.5: representation of two different phases of the machining process of a component using a redundant additive manufacturing device

4.4.3 The two pillars of the integrated design

As it had been already stated, the integrated design is the definition of the best structure and control for a device. According to the presented framework, the structure of the mechanism can be modified introducing a redundant degree of freedom. Moreover, in order to improve the performances of the mechanism, it is necessary to set up an optimization problem: it is possible because the introduction of the redundancy in the kinematic chain allows the definition of a secondary task, and, consequently, of an objective function. Actually, the modify of the morphology of the mechanism, and the control based on the solution of the consequent indeterminacy are an instance of the two pillars of the general methodology introduced in chapter 1.

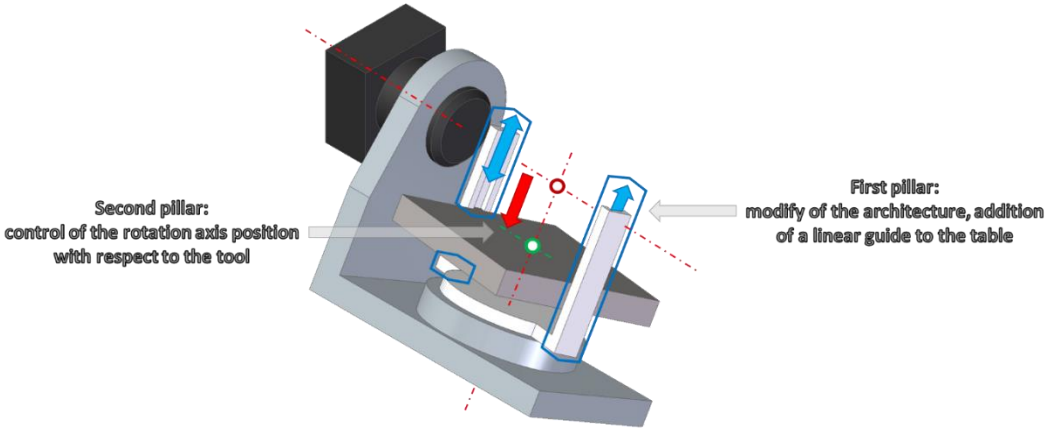
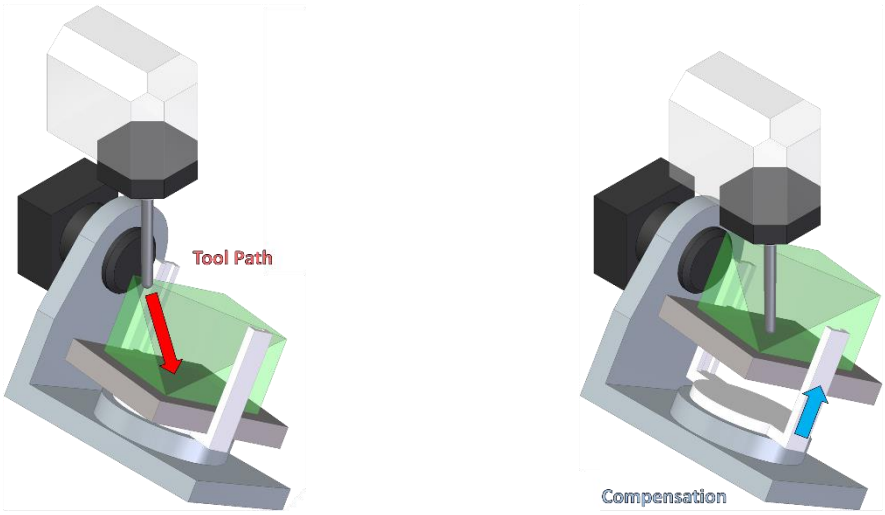


Figure 4.6: implementation of the two pillars of integrated design for the design of a novel CNC turntable

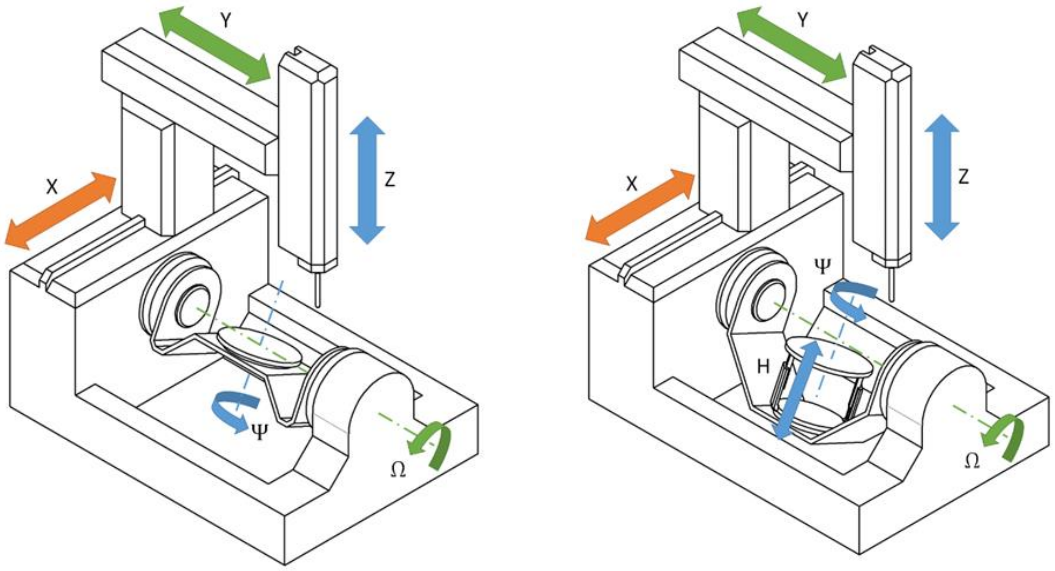
In the case of the re-design of the 5-axis milling machine, or 3D printer, or hybrid machine tool, the support table for the workpiece is modified adding a further linear actuator, as it is shown in figure 4.6.



(a) *(b)*

Figure 4.7: control of the redundant linear axis during the milling process

As it had been shown in the previous subsections, as a consequence of the introduction of a redundant degree of freedom, there are infinite configuration of the machine that realize the prescribed position and pose of the tool with to respect the workpiece. For this reason, it is necessary to introduce in the control logic a criterion for the choice of a preferred configuration, and this is done by the solution of an optimization problem.



(a) *(b)*

Figure 4.8: layout of the RRTTT machine tool (a), and layout of the TRRTTT machine tool (b)

As a first example of objective function, it is possible to impose the minimization of the sensitivity of the positioning error of the tool position to the angular errors of the rotary actuator of the table. In other words, as it is shown in figure 4.7, it is possible to realize a control that, while the tool moves along it prescribed trajectory with respect to the reference framework attached to the workpiece, minimizes the distance between the tooltip and the rotation axis, consequently minimizing the position errors. In the next section it will be provided the theoretical background of the implementation of the methodology to this specific case study, and, due to the use of a virtual machine model, some preliminary results will be presented.

4.4.4 Theoretical aspects and verification

A 5-axis milling machine is composed by five links, six considering the ground link, which are connected by five kinematic couple, usually three prismatic guides, and two rotary joints. The kinematic chain is serial, meaning that there are not closed chains.

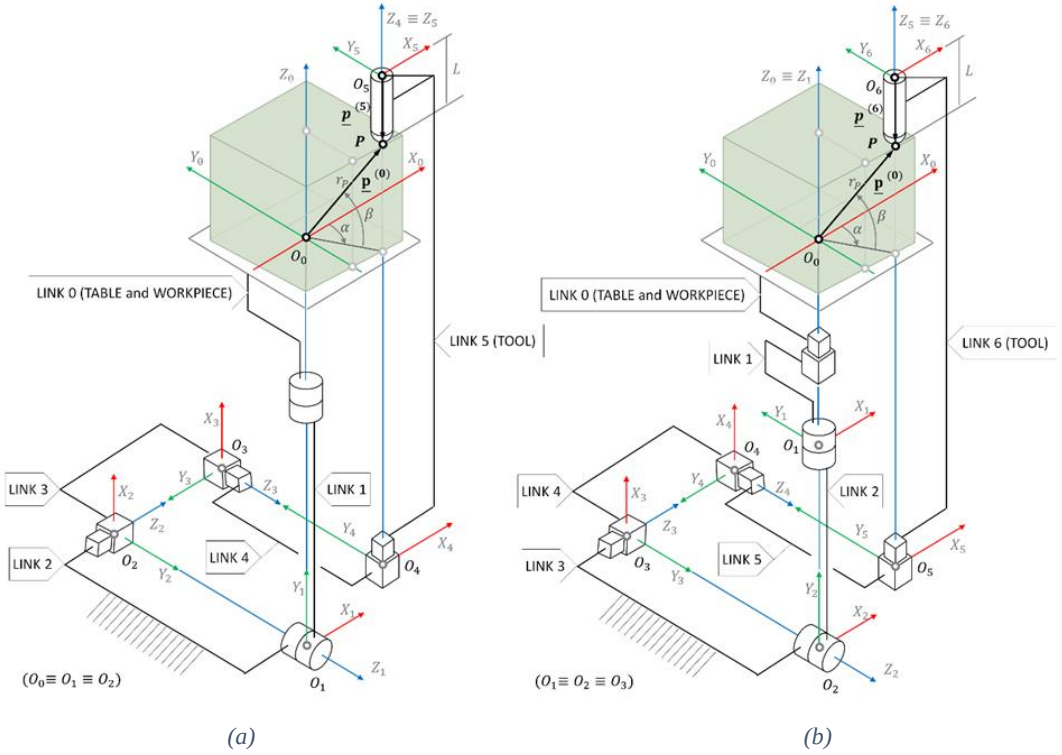


Figure 4.9: kinematic chain of the RRTTT machine tool (a), and kinematic chain of the TRRTTT machine tool (b)

As a convention, different machines are designated by indicating the order of the kinematic couples, starting from the joint directly connected to the workpiece. In the present research, the original mechanism is indicated as the RRTTT 5 axis machine tool, depicted in the left figure 3.8. According to the theoretical background provided in the previous section, the RRTTT 5 axis machine tool is a non-redundant mechanism, which means that the inverse kinematic problem of positioning the tool with to respect the workpiece has a finite number of solutions.

Let us introduce a redundant, actuated degree of freedom in the kinematic chain, as it is depicted in right figure 3.8; this is a 6 axis TRRTTT machine tool, and its inverse kinematic problem is no more determined, and it accept an infinite number of solutions. In order to formulate a new determined problem, it is necessary to integrate a new criterium, or, in other words, an objective function to optimize.

As a first step, it is necessary to write the closure equations for the 6 axis TRRTTT milling machine. This is done using the standard procedure provided by the Denavit and Hartemberg convention: as a first step, there will be derived the closure equations for the

RRTTT 5 axis machine tool, and then, the equation for the kinematic chain for the TRRTTT 6 axis machine tool.

Closure equations for the RRTTT milling machine

The Denavit and Hartemberg convention allows to express the position of a point with to respect a reference system attached to a certain link of the mechanism, in function of the position of the same point with to respect the reference system of another link. Let $O_0x_0y_0z_0$ be the reference system attached to the workpiece, and $O_5x_5y_5z_5$ the reference system attached to the tool; furthermore, if $\underline{p}^{(0)}$ is the vector of the coordinates of a point P with respect to the reference coordinate system $O_0x_0y_0z_0$, and $\underline{p}^{(5)}$ is the vector of the coordinates of a point P with respect to the reference coordinate system $O_5x_5y_5z_5$, it is possible to write:

$$\underline{p}^{(5)} = \underline{A}_0^5 \cdot \underline{p}^{(0)} \tag{4.23}$$

where $\underline{A}_0^5$ is global homogenous transformation matrix defined by the following expression:

$$\underline{A}_5^0 = \underline{A}_1^0 \cdot \underline{A}_2^1 \cdot \underline{A}_3^2 \cdot \underline{A}_4^3 \cdot \underline{A}_5^4 \tag{4.24}$$

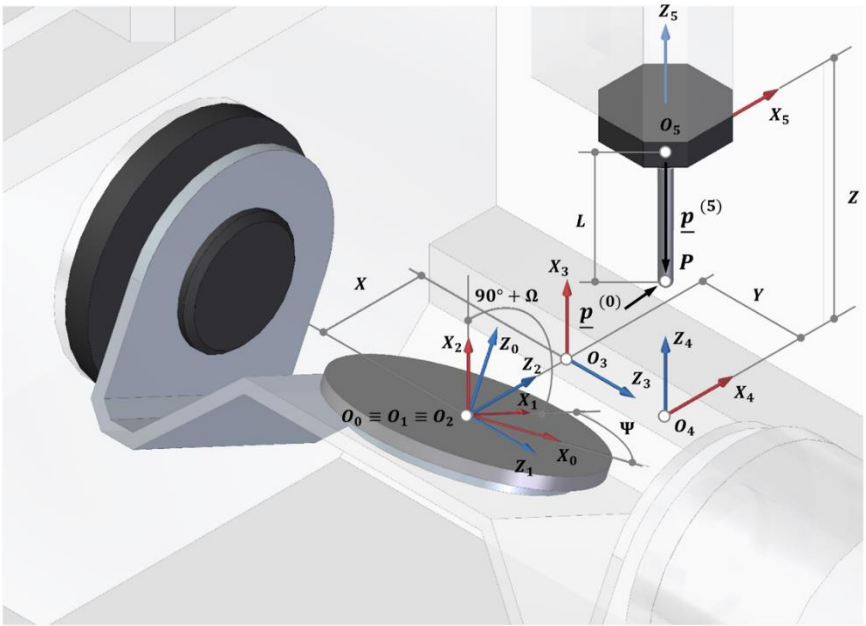


Figure 4.10: set of reference systems and Lagrange variables for the RRTTT machine tool

All the factors of the previous expression are homogenous transformation matrix for a single link of the kinematic chain of the mechanism, and, in general, they have the following form:

$$\underline{\underline{A}}_i^{i-1}(\,d_i\,,\theta_i\,,a_i\,,\alpha_i\,) = \begin{vmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{4.25}$$

The complete set of reference systems and Lagrange variables competing all the links of the mechanism are depicted in figure 4.10 and the complete set of parameters are reported in table 1

	$\boldsymbol{d_i}$	$\boldsymbol{\theta_i}$	$\boldsymbol{a_i}$	$\boldsymbol{\alpha_i}$
1	0	Ψ	0	90°
2	0	$90° + \Omega$	0	90°
3	X	0	0	$-90°$
4	Y	$-90°$	0	$-90°$
5	Z	0	0	0

Table 4.1

The final global transformation matrix can be expressed in function of the Lagrange variables of the mechanism Ψ, Ω, X, Y, Z:

$$\underline{\underline{A}}_5^0 = \begin{vmatrix} c\Psi c\Omega & -s\Psi & -c\Psi s\Omega & (Xc\Psi c\Omega + Ys\Psi - Zc\Psi s\Omega) \\ s\Psi c\Omega & c\Psi & -s\Psi s\Omega & (Xs\Psi c\Omega - Yc\Psi - Zs\Psi s\Omega) \\ s\Omega & 0 & c\Omega & (Xs\Omega + Zc\Omega) \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{4.26}$$

Ψ, Ω represent angular coordinates, and X, Y, Z are linear coordinates. According to [11], the kinematic chain equations read as follows:

$$\begin{cases} x_p = Lc\Psi s\Omega + Xc\Psi c\Omega + Ys\Psi \\ y_p = Ls\Psi s\Omega + Xs\Psi c\Omega - Yc\Psi \\ z_p = Lc\Omega + H + Xs\Omega \\ \underline{k}^{(5)} \cdot \underline{i}^{(0)} = -c\Psi s\Omega \\ \underline{k}^{(5)} \cdot \underline{k}^{(0)} = c\Omega \end{cases} \tag{4.27}$$

and, consequently, the Jacobian may be written as follows:

$$\underline{\underline{J}}_{5x5}^{(0)} = \begin{vmatrix} [-Xs\Psi c\Omega + Yc\Psi - Ls\theta_1 s\Omega] & [-Xc\Psi s\Omega + Lc\Psi c\Omega] & c\Psi c\Omega & s\Psi & -c\Psi s\Omega \\ [Xc\Psi c\Omega + Ys\Psi + Lc\theta_1 s\Omega] & [-Xs\Psi s\Omega + Ls\Psi c\Omega] & s\Psi c\Omega & c\Psi & -s\Psi s\Omega \\ 0 & [-Xc\Omega + Ls\Omega] & s\Omega & 0 & c\Omega \\ s\Psi s\Omega & -c\Psi c\Omega & 0 & 0 & 0 \\ 0 & -s\Omega & 0 & 0 & 0 \end{vmatrix} \quad (4.28)$$

Closure equations for the TRRTTT milling machine

The kinematic closure equations for the TRRTTT 6 axis mailing machine are derived in the same fashion of the RRTTT 5 axis milling machine model: $O_0x_0y_0z_0$ is the reference system attached to the workpiece, $O_6x_6y_6z_6$ is the reference system attached to the tool, $\underline{p}^{(0)}$ is the vector of the coordinates of a point P with respect to the reference coordinate system $O_0x_0y_0z_0$, and $\underline{p}^{(6)}$ is the vector of the coordinates of a point P with respect to the reference coordinate system $O_6x_6y_6z_6$.

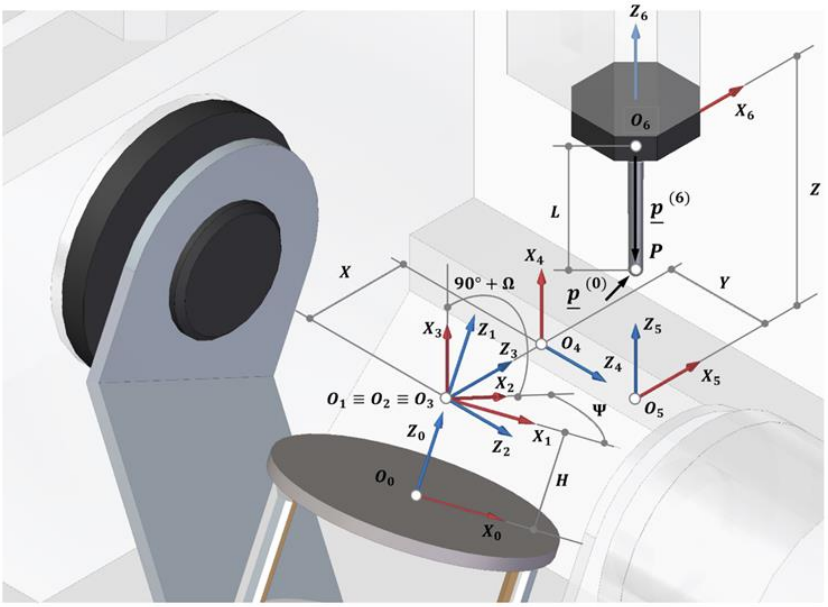


Figure 4.11: set of reference systems and Lagrange variables for the TRRTTT machine tool

Furthermore, as reported in figure 4.9, the position vector for the point P with to respect the coordinate system $O_0x_0y_0z_0$ is a fuction of the parameters r_P, α, β, and it is related to the vector position for the point P with to respect the frame $O_6x_6y_6z_6$ by the following relation

$$\underline{p}^{(6)} = \begin{vmatrix} 0 \\ 0 \\ -L \end{vmatrix} = \underline{\underline{A}}_0^6 \cdot \underline{p}^{(0)} = \underline{\underline{A}}_0^6 \cdot \begin{vmatrix} r_P c\alpha c\beta \\ r_P s\alpha c\beta \\ r_P s\beta \end{vmatrix} \tag{4.29}$$

where $\underline{\underline{A}}_0^6$ is the homogenous transformation matrix that allows the passage from the representations of the point P in the $O_0x_0y_0z_0$, $\underline{p}^{(0)}$, to the representations of the point P in the $O_6x_6y_6z_6$, $\underline{p}^{(6)}$.

	$\boldsymbol{d_i}$	$\boldsymbol{\theta_i}$	$\boldsymbol{a_i}$	$\boldsymbol{\alpha_i}$
0	H	0	0	0
1	0	Ψ	0	90°
2	0	$90° + \Omega$	0	90°
3	X	0	0	$-90°$
4	Y	$-90°$	0	$-90°$
5	Z	0	0	0

Table 4.2

Similarly to what was done previously, the complete set of reference systems and Lagrange variables for the TRRTTT 6 axis milling machine competing all the links of the mechanism are depicted in figure 4.11, and the complete set of parameters are reported in table 2. Consequently, the homogenous coordinate transformation matrix is:

$$\underline{\underline{A}}_6^0 = \underline{\underline{A}}_1^0 \cdot \underline{\underline{A}}_2^1 \cdot \underline{\underline{A}}_3^2 \cdot \underline{\underline{A}}_4^3 \cdot \underline{\underline{A}}_5^4 \cdot \underline{\underline{A}}_6^5 = \begin{vmatrix} c\Psi c\Omega & -s\Psi & -c\Psi s\Omega & (Xc\Psi c\Omega + Ys\Psi - Zc\Psi s\Omega) \\ s\Psi c\Omega & c\Psi & -s\Psi s\Omega & (Xs\Psi c\Omega - Yc\Psi - Zs\Psi s\Omega) \\ s\Omega & 0 & c\Omega & (H + Xs\Omega + Zc\Omega) \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{4.30}$$

Ψ, Ω represent angular coordinates, and X, Y, Z are linear coordinates, and and H is the variable parameter introduced with the redundant kinematic couple. The kinematic chain equations read as follows

$$\begin{cases} x_p = Lc\Psi s\Omega + Xc\Psi c\Omega + Ys\Psi - Zc\Psi s\Omega \\ y_p = Ls\Psi s\Omega + Xs\Psi c\Omega - Yc\Psi - Zs\Psi s\Omega \\ z_p = Lc\Omega + H + Xs\Omega + Zc\Omega \\ \underline{k}^{(6)} \cdot \underline{i}^{(0)} = -c\Psi s\Omega \\ \underline{k}^{(6)} \cdot \underline{k}^{(0)} = c\Omega \end{cases} \tag{4.31}$$

or, equivalently, with respect to the reference coordinate system $O_6x_6y_6z_6$, which is attached to the tool, the final kinematic chain equations read as follows

$$\begin{cases} 0 = Lc\Psi s\Omega + Xc\Psi c\Omega + Ys\Psi - Zc\Psi s\Omega \\ 0 = Ls\Psi s\Omega + Xs\Psi c\Omega - Yc\Psi - Zs\Psi s\Omega \\ -L = Lc\Omega + H + Xs\theta_2 + Zc\Omega \\ \underline{k}^{(6)} \cdot \underline{i}^{(0)} = -c\Psi s\Omega \\ \underline{k}^{(6)} \cdot \underline{k}^{(0)} = c\Omega \end{cases} \tag{4.32}$$

An interesting remark is that, due to the fact the slope of the tool with respect to the work piece is only referred to the coordinates Ψ and Ω of the rotational axes, only the equations competing the translations are changed.

Finally, the Jacobian may be written as follows

$$\underline{\underline{J}}_{6x5}^{(0)} = \begin{vmatrix} 0 & [-Xs\Psi c\Omega + Yc\Psi - (L-Z)s\theta_1 s\Omega] & [-Xc\Psi s\Omega + (L-Z)c\Psi c\Omega] & c\Psi c\Omega & s\Psi & -c\Psi s\Omega \\ 0 & [Xc\Psi c\Omega + Ys\Psi + (L-Z)c\theta_1 s\Omega] & [-Xs\Psi s\Omega + (L-Z)s\Psi c\Omega] & s\Psi c\Omega & c\Psi & -s\Psi s\Omega \\ 1 & 0 & [-Xc\Omega + (L+Z)s\Omega] & s\Omega & 0 & c\Omega \\ 0 & s\Psi s\Omega & -c\Psi c\Omega & 0 & 0 & 0 \\ 0 & 0 & -s\Omega & 0 & 0 & 0 \end{vmatrix} \tag{4.33}$$

Optimal configuration of the TRRTTT machine

Since the TRRTTT 6 axis milling machine is a redundant mechanism to the task of the correct positioning of the tool with respect to the workpiece, it is necessary to define a criterion to eliminate the indeterminacy introduced by the redundant degree of freedom. It had been already stated that this is done by the implementation of an optimization problem in the control logic. As it had been already shown in chapter 3, a standard form for an optimization problem reads as follows

$$\begin{cases} min. \quad U(\underline{x}) \\ s.t. \quad V_j(\underline{x}) = 0 \qquad j = 1, \dots, J \\ \quad\quad W_k(\underline{x}) \geq 0 \qquad k = 1, \dots, K \\ \quad\quad x_i^{(low)} \leq x_i < x_i^{(high)} \qquad i = 1, \dots, N \end{cases} \tag{4.34}$$

$U(\underline{x})$ is the objective function, the mathematical expression of the criterion adopted for the elimination of the mathematical indeterminacy of the inverse kinematic problem for the redundant mechanism, $\underline{x}$ is the vector containing the state variables of the problem: in the previous chapter, this vector had a large number of components, corresponding to the pseudo-densities of the finite elements of the continuum model. In this case, the state variables are the Lagrange variables describing the configuration of the mechanism (control variables); for the TRRTTT 6 axis milling machine these quantities are H, Ψ, Ω, X, Y, and Z, representing the axis positions of the actuators of the mechanism.

More specifically, here will be presented two different objective functions: the first one is the norm of the positional errors due to the wrong actual position of rotational joints,

already introduced in the previous subsection, and the second one is related to the torques provided by the rotational motors during the milling process. It is important to highlight that these two particular objective functions represent two performances indexes for the mechanism.

The equations $V_j(\underline{x})$ are the equality constraints of the optimization problem, and, for the optimization of the control of a mechanism composed by rigid links, are the closure equations of the kinematic chain: in fact, this are expressions that must satisfied because represent the kinematic characterization of the mechanism.

The inequality constraints are not taken in account here because they have no practical use: differently from what have been done in the chapter 3, where the optimization problem has many control variables, and it was not possible to obtain a closed form solution (for this reason gradient based methods are involved to obtain a solution), here it is possible to have a direct solution formulation; for this reason, the only inequality condition present in this problem, which are the limit on the run of the linear guides, can be directly checked by substitution in the solution equation.

The efficiency of the methodology will now be demonstrated, according to the previous defined objective functions, considering the optimal machine configuration for 2D machining problems. The 2D machining is derived by the general 3D formulation by imposing the following values for the control variables: $\theta_1 = 0,\ \alpha = 0,\ Y = 0$. As a further condition, the origin of the reference system $O_6 x_6 y_6 z_6$ is set coincident with the tool centre, which implies that $L = 0$ as well. Under these hypothesis, the kinematic closure equations become:

$$\begin{cases} 0 = r_P c\beta c\Omega + r_P s\beta s\Omega - Hs\Omega - X \\ 0 = 0 \\ 0 = -r_P c\alpha c\beta s\Omega + r_P s\beta c\Omega - Hc\Omega - Z \\ \underline{k}^{(6)} \cdot \underline{i}^{(0)} = -s\Omega \\ \underline{k}^{(6)} \cdot \underline{k}^{(0)} = c\Omega \end{cases} \Longrightarrow \tag{4.35}$$

$$\Longrightarrow \begin{cases} 0 = r_P \cos(\beta - \Omega) - Hs\Omega - X \\ 0 = r_P \sin(\beta - \Omega) - Hc\Omega - Z \\ \mu = \Omega + \dfrac{\pi}{2} \end{cases}$$

where μ is the angle between the x_0 axis attached to the work piece, and the axis of the tool z_6. Moreover, it is possible to provide the Jacobian matrix of the system according

to the 2D formulation of the problem, and the consequent change of the dimensions of the matrix itself:

$$\underline{\underline{J}}_{3x4}^{(6)} = \begin{vmatrix} -s\Omega & r_P\sin(\beta-\Omega) - Hc\Omega & -1 & 0 \\ -c\Omega & -r_P\cos(\beta-\Omega) + Hs\Omega & 0 & -1 \\ 0 & 1 & 0 & 0 \end{vmatrix} \tag{4.36}$$

Position error minimization

Let us start considering the minimization of the position error norm. Chen [77] highlighted that position errors have a high sensitivity to angular position of rotation motors and angular errors. In order to evaluate the entity of its contribution to volumetric errors, it is possible to write infinitesimal variation of the positional parameters in function of the infinitesimal variation of the axis coordinates

$$\begin{cases} dx_P^{(6)} &= -s\Omega \cdot dH + (r_P\sin(\beta-\Omega) - Hc\Omega)\cdot d\Omega - dX \\ dz_P^{(6)} &= -c\Omega \cdot dH + (-r_P\cos(\beta-\Omega) + Hs\Omega)\cdot d\Omega - dZ \end{cases} \tag{4.37}$$

The objective function is the norm of the vector representing the positioning error, and, consequently, the second optimization problem may be written in the following form

$$\begin{array}{llllllll} min. & \widehat{U}_1(H,\Omega,X,Z,d\Omega) & = & \left\|\left(dx_P^{(6)}, dz_P^{(6)}\right)\right\| & = & \sqrt{\left(dx_P^{(6)}\right)^2 + \left(dz_P^{(6)}\right)^2} & = & ds_P^{(6)} \\ s.t. & V_1(H,\Omega,X,Z,d\Omega) & & & = & r_P\cos(\beta-\Omega) - Hs\Omega - X & = & 0 \\ & V_2(H,\Omega,X,Z,d\Omega) & & & = & r_P\sin(\beta-\Omega) - Hc\Omega - Z & = & 0 \\ & V_3(H,\Omega,X,Z,d\Omega) & & & = & \mu - \left(\Omega + \frac{\pi}{2}\right) & = & 0 \end{array} \tag{4.38}$$

The equivalence of the notations $\widehat{U}_1(H,\Omega,X,Z,d\Omega) = \widehat{U}_1(H,\Omega,d\Omega)$ and $V_i(H,\theta_2,X,Z,d\theta_2) = V_i(H,\theta_2,X,Z)$ depends by the fact that the kinematic closure equations do not depend by the infinitesimal rotation $d\Omega$. Furthermore, minimizing the objective function $\widehat{U}_2$ is equivalent to minimizing its square divided by the quantity $(d\Omega)^2$, which is a positive quantity. For this reason, it is possible define a new equivalent optimization problem:

$$\begin{array}{llllll} min. & U_1(H,\Omega,d\Omega) & = & \dfrac{\left(dx_P^{(6)}\right)^2 + \left(dz_P^{(6)}\right)^2}{(d\Omega)^2} & = & \dfrac{\left(ds_P^{(6)}\right)^2}{(d\Omega)^2} \\ s.t. & V_1(H,\Omega,X,Z) & = & r_P\cos(\beta-\Omega) - Hs\Omega - X & = & 0 \\ & V_2(H,\Omega,X,Z) & = & r_P\sin(\beta-\Omega) - Hc\Omega - Z & = & 0 \\ & V_3(H,\Omega,X,Z) & = & \mu - \left(\Omega + \dfrac{\pi}{2}\right) & = & 0 \end{array} \tag{4.39}$$

This is a convex constrained optimization problem, and this means that exist a single global minimum. The expression for the global minimum may be directly derived, and it

is $U_{1opt} = {r_P}^2 \cos^2\beta$, and it corresponds to the value of the lagrangian variable $H_{opt1} = r_P sin(\beta)$. According to this condition, it is possible to determine the optimal configuration of the mechanism, which is depicted in figure: it is interesting to highlight that the point P lays on the axis $x_1 \equiv x_2$, and this actual means that the distance between the point P and the point O_1 is minimum.

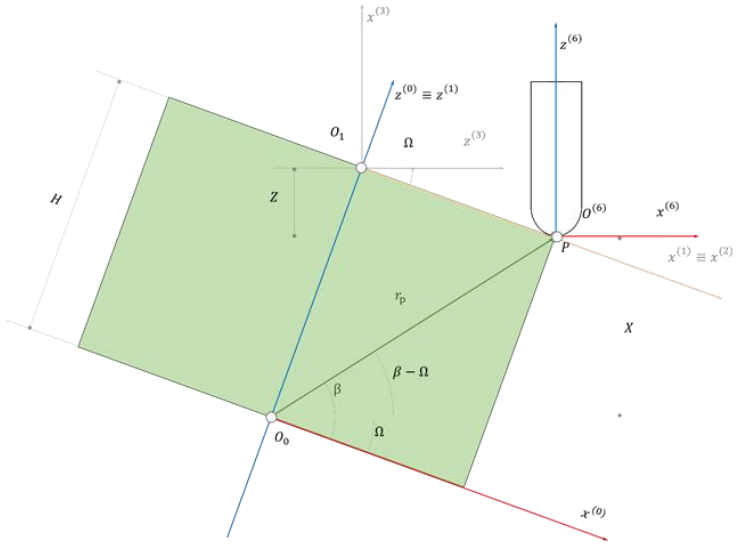


Figure 4.12: graphical representation of the optimality condition for the position error minimization

Actuator torque minimization

The minimization of the position error is not the only feasible optimization criterium feasible to be integrated in the control of the TRRTTT 6 axis machine tool. Let us consider again a 2D milling process, and define $F_x^{(6)}$, $F_z^{(6)}$, $M_y^{(6)}$ as the components of the force and the momentum applied by the tool on the work piece. In the previous sections it had been stated that, by the means of the Jacobian matrix, it is possible to solve the static of the system due to the kinetic/static duality, being

$$\begin{vmatrix} T_H \\ C_\Omega \\ T_X \\ T_Z \end{vmatrix} = \underline{\underline{G}}_{4x3}^{(6)} \cdot \begin{vmatrix} F_x^{(6)} \\ F_z^{(6)} \\ M_y^{(6)} \end{vmatrix}$$

$$\Rightarrow \begin{cases} T_H = -\left(s\Omega F_x^{(6)} + c\Omega F_z^{(6)}\right) \\ C_\Omega = (r_P \sin(\beta - \Omega) - Hc\theta_2) \cdot F_x^{(6)} - (r_P \cos(\beta - \Omega) - Hs\Omega) \cdot F_z^{(6)} + M_y^{(6)} \\ T_X = -F_x^{(6)} \\ T_Z = -F_z^{(6)} \end{cases} \quad (4.40)$$

Where, actually, the matrix $\underline{\underline{G}}_{4x3}^{(6)} = \left(\underline{\underline{J}}_{3x4}^{(6)}\right)^T$ is the transpose of the Jacobian matrix.

T_H, C_Ω, T_X, and T_Z are the forces and torques provided by the actuator in order to ensure the equilibrium if the mechanism subject to the external loads, and correspond to the lagrange variables H, Ω, X, and Z respectively.

According to the constraints of the optimization problem (the kinematic chain equations) it is possible to notice that T_X and T_Z depends only by $F_x^{(6)}$ and $F_z^{(6)}$, and, for this reason, are prescribed and may not vary; furthermore, if the milling parameters are provided, T_H can assume only one value, because Ω is the only parameter available for the definition of the pose of the tool with to respect the workpiece.

On the contrary, even if pose and position are given, it is not sufficient to completely define C_Ω, because there are infinite configurations that can fulfil the position requirements. This means that the value of the couple C_2 can be adopted as objective function

$$\begin{array}{llcccl} min. & U_2(H,\Omega,X,Z) & = & \left|(r_P\sin(\beta-\Omega)-Hc\Omega)\cdot F_x^{(6)}+(r_P\sin(\beta-\Omega)-Hs\Omega)\cdot F_z^{(6)}\right| & = & |C_\Omega| \\ s.t. & V_1(H,\Omega,X,Z) & = & r_P\cos(\beta-\Omega)-Hs\Omega-X & = & 0 \\ & V_2(H,\Omega,X,Z) & = & r_P\sin(\beta-\Omega)-Hc\Omega-Z & = & 0 \\ & V_3(H,\Omega,X,Z) & = & \mu-\left(\Omega+\frac{\pi}{2}\right) & = & 0 \end{array} \quad (4.41)$$

In the 2D case, the objective function can be nullified, reaching a global minimum, imposing the following condition for H:

$$\frac{F_z^{(6)}}{F_x^{(6)}} = \frac{\left(r_P\sin(\beta-\Omega)-H_{opt2}c\Omega\right)}{\left(r_P\cos(\beta-\Omega)-H_{opt2}s\Omega\right)} \quad (4.42)$$

Figure 4.13 depicts the graphical representation of the optimality condition: as it may be expected, the ideal configuration for the machine shows the line of action of the force passing through the point of instant rotation. This condition, if theoretically may be always met, in practice cannot because the parameter H may only have values in accordance with the physical limits of the mechanical linear guide.

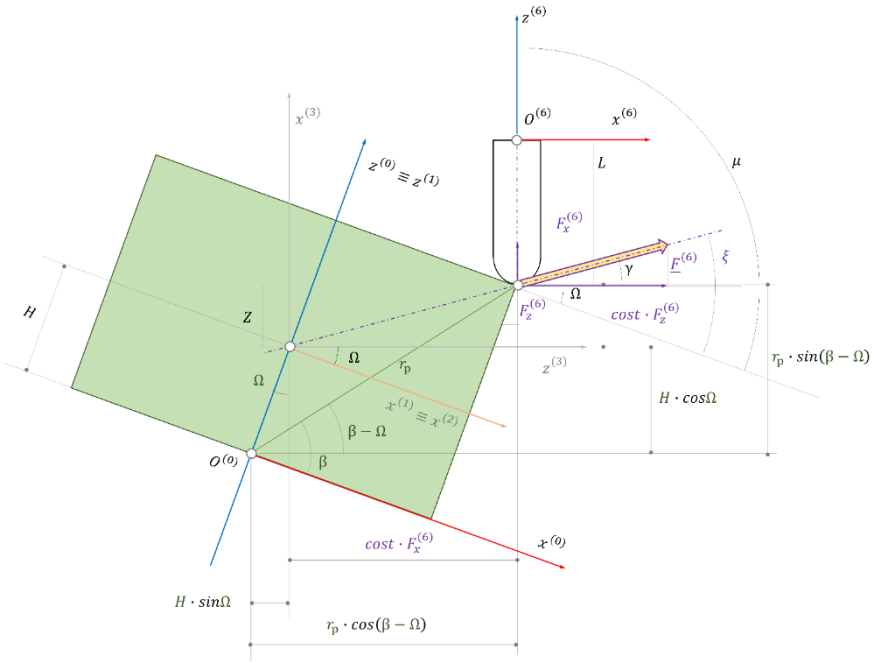
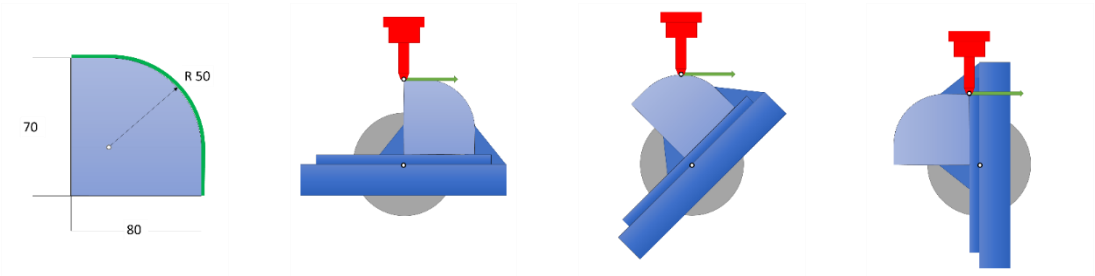


Figure 4.13: graphical representation of the optimality condition for the actuator torque minimization

Virtual prototype and results: positioning error minimization

Let us consider the milling process depicted in figure 4.14 as a case study in order to evaluate the effectiveness of the proposed approach. Since a physical prototype of the TRRTTT 6 axis milling machine does not exist yet, a simulation of the milling process by the mean of a virtual model is a convenient approach in order to obtain preliminary evaluations. Moreover, in a further stage of development of the machine, the virtual model may be easily adapted for the implementation of the control logic.



(a) *(b)* *(c)* *(d)*

Figure 4.14: tool path for the case study machining procedure

More in detail, as it is depicted in figure 4.14, the path of the trajectory is composed by a first linear segment with 30 mm length, a circular arc of 50 mm radius, and a second linear segment of 20 mm. During the machining process, the force on the work piece is always tangent to the trajectory.

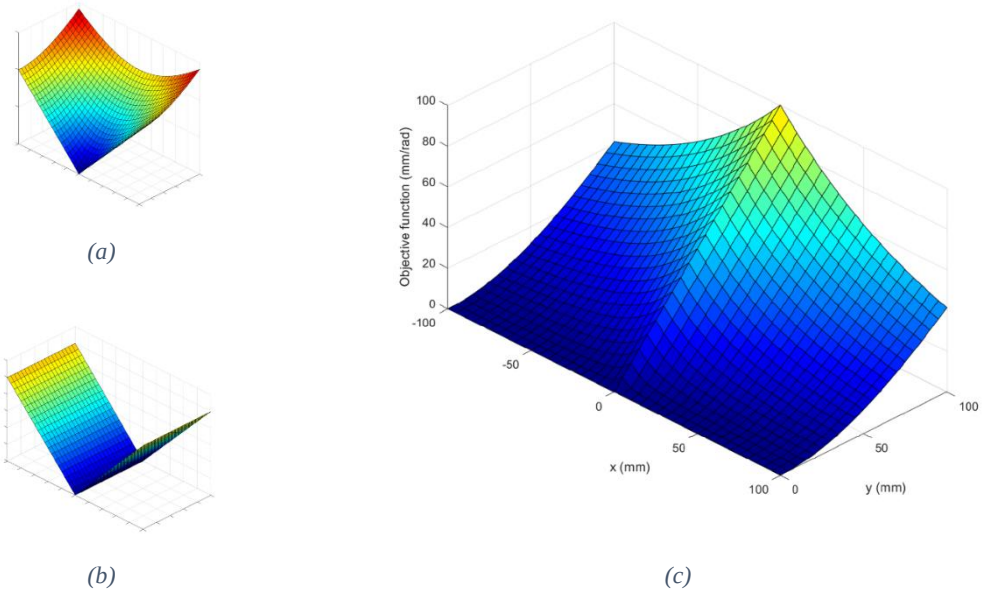
The simulations of the optimization algorithm implementation for both positioning error minimization and motor torque minimization have been caried out by the mean of MATLAB scripts: all the procedures and functions have been written using the basic features of the computational environment, and the output has been produced using the standard graphical tools.

Let us start with the positioning error minimization: recalling the expression of the objective function U_1

$$U_1(H, \Omega, X, Z) = H^2 - 2Hr_P sin(\beta) + {r_P}^2 \tag{4.43}$$

figure 4.15(a) shows the mapping of the position error norm in the case of the standard 5 axis RRTTT milling machine on the workspace for the tool centre for a work piece of 200x100 mm. On the other hand, figure 4.15(b) shows the mapping of the position error norm in the case of the 6 axis TRRTTT milling machine, for the same workpiece, if it satisfied the optimal configuration condition:

$$H_{opt1} = r_P sin(\beta) \tag{4.44}$$



(a)

(b)

(c)

Figure 4.15: mapping of the position error for the RRTTT milling machine (a), mapping of the position error for the TRRTTT milling machine (c), error improvement (c)

As an interesting remark, it may be noticed that, even adopting the work piece table compensation, it is possible to nullify the errors in correspondence of the points laying on the z_0 axis; this depends by the fact that the only points of the work piece which can be coincident with the rotation centre are the points on the z_0 axis.

Consequently, figure 4.15(c) shows the difference between the position errors in the two cases, highlighting the fact that it is always a positive or, at least, a null improvement.

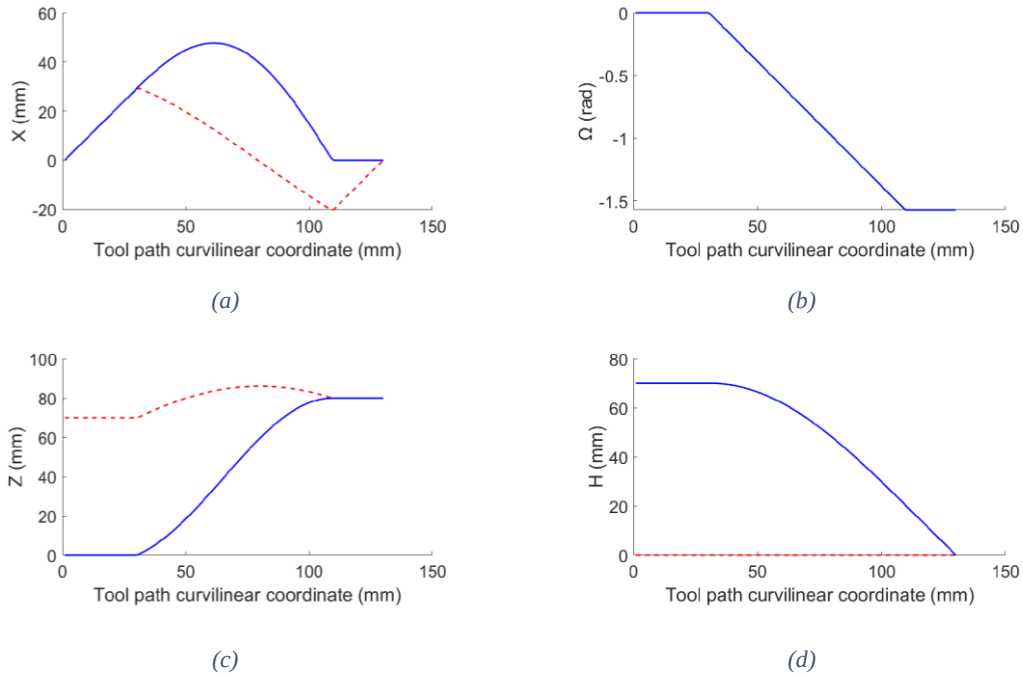


Figure 4.16: H, Ω, X, and Z in function of the curvilinear coordinate for positioning error minimization

Figure 4.16 shows the evolution of the Lagrange variables H, Ω, X, and Z, as functions of the toolpath curvilinear coordinate in the case of the milling process depicted in figure 4.14, using an RRTTT 5 axis milling machine (dashed red line), and in the case of a TRRTTT 6 axis milling machine (continuous blue line).

The reported data completely describe the configuration of the two machines during the milling process. Consequently, it is possible to determine the evolution of the objective function during the machining of the profile of the work piece. It can be noticed that the values are different, except for the variable Ω which, in both machines, is the only variable involved in the determination of the slope of the work piece respect to the tool. Moreover,

for the RRTTT 5 axis machine tool the variable H is identically equal to zero, because it misses the redundant axis.

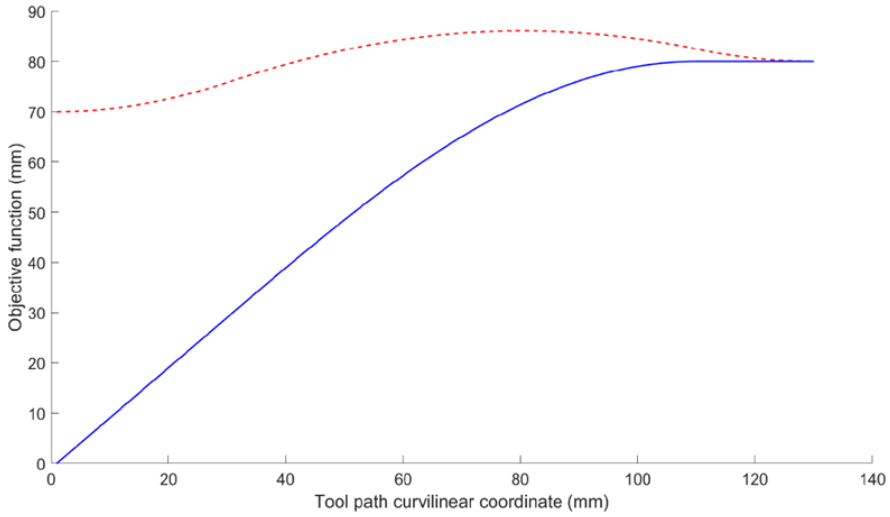


Figure 4.17: evolution of U_1 with respect to the curvilinear coordinate of the tool path

Moreover, figure 4.17 shows the evolution of the objective function with respect to of the toolpath curvilinear coordinate in the case of the milling process depicted in figure 4.14, using an RRTTT 5 axis milling machine (dashed red line), and in the case of a TRRTTT 6 axis milling machine (continuous blue line). It can be noticed that the TRRTTT 6 axis milling machine has always better or equal performances compared to the RRTTT 5 axis milling machine.

Virtual prototype and results: actuator torque minimization

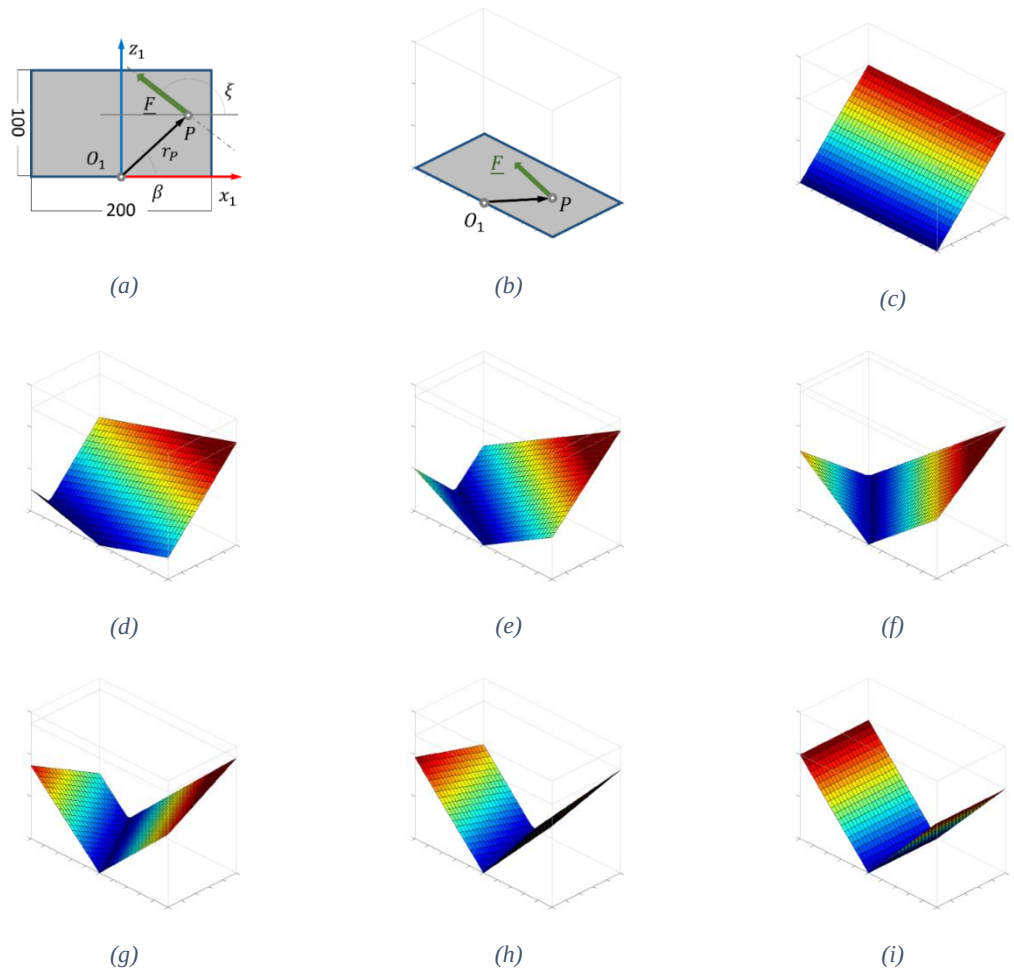
The second case study is the minimization of the torque provided by the rotary motor in order to fulfil the equilibrium condition of the machine during the milling process, and this is reasonable because the velocities of the tool and work piece are limited, and the inertia forces can be ignored.

In this hypothesis, the objective function U_2 may be written as:

$$U_2(H, \Omega, X, Z) = F \cdot (r_P \, sin(\beta - \xi) + H cos\xi) \tag{4.45}$$

Figure 4.18 shows the mapping of U_2 on the workspace for the tool centre for a work piece of 200x100 mm, considering different slopes ξ of the force $\underline{F}$. Unlike the previous case, here the two distinct cases for the RRTTT 5 axis milling machine and the TRRTTT

6 axis milling machine are not reported, because they are basically equivalent, despite a shifting along the z_1 axis, due to the presence of the redundant variable H. In the best conditions, the torque required to the rotation motor is zero in the points laying on lines having the same inclination of the force $\underline{F}$, and passing through the rotation point O_1.



(a) *(b)* *(c)*

(d) *(e)* *(f)*

(g) *(h)* *(i)*

Figure 4.18: mapping of U_2 on the workspace for the tool centre for a work piece of 200x100 mm (a) and (b) for ξ=0° (c), ξ=15° (c), ξ=30° (c), ξ=45° (c), ξ=60° (c), ξ=75° (c), ξ=90° (c)

In the case of the milling process described in figure 4.14, it may be noticed that the interaction force between the tool and the workpiece is always tangent to the trajectory. Consequently, the angle γ is always equal to 0, and, taking in account that and remembering that ξ = γ + Ω, it is possible to write:

$$U_2(H, \Omega, X, Z) = F \cdot (r_P \, sin(\beta - \Omega) + H cos\Omega) \tag{4.46}$$

The value of the Lagrange variable H satisfying the minimum required torque optimality criteria can be directly calculated by the minimization of the objective function

$$H_{opt2} = r_P \cdot \frac{\sin(\beta - \xi)}{cos(\xi)} = r_P \cdot \frac{\sin(\beta - \Omega)}{cos(\Omega)} \tag{4.47}$$

Which is the ideal configuration of the machine in correspondence of every point of the trajectory of the centre of the tool.

Figure 4.19 shows the evolution of the Lagrange variables H, Ω, X, and Z, as functions of the toolpath curvilinear coordinate in the case of the milling process depicted in figure 4.14, using an RRTTT 5 axis milling machine (dashed red line), and in the case of a TRRTTT 6 axis milling machine (continuous blue line).

Moreover, figure4.20 shows the evolution of the objective function with respect to of the toolpath curvilinear coordinate, using an RRTTT 5 axis milling machine (dashed red line), and in the case of a TRRTTT 6 axis milling machine (continuous blue line).

It can be noticed that the TRRTTT 6 axis milling machine has always better or equal performances compared to the RRTTT 5 axis milling machine.

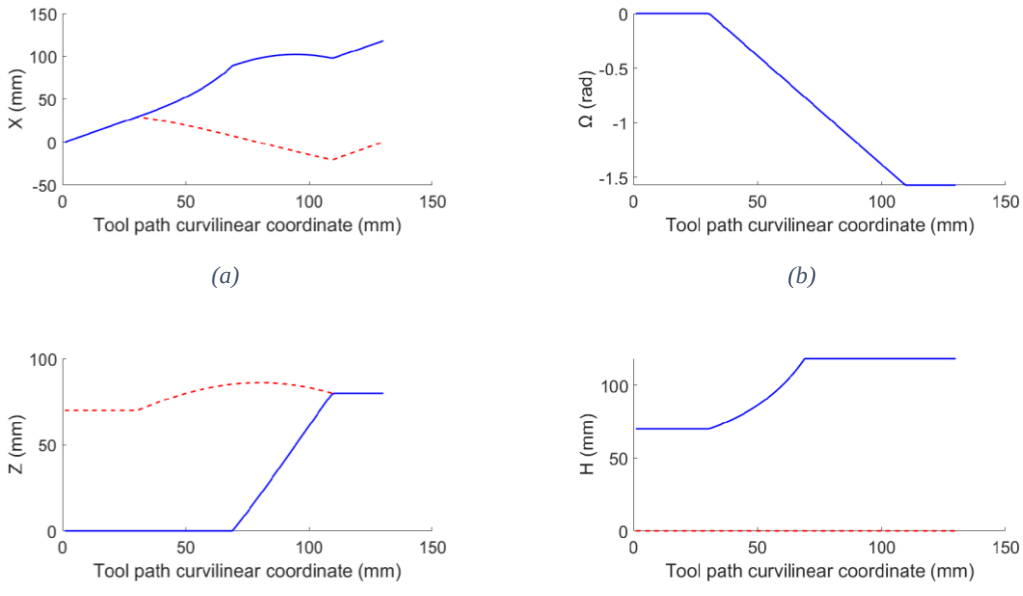


Figure 4.19: H, Ω, X, and Z in function of the curvilinear coordinate for motor torque minimization

The introduction of the redundant degree of freedom allows the limitation of the torque required to the rotational axis. This appends for almost every point of the trajectory, except the final part. This may be explained considering the fact that, even if in theory the variation of the variable *H* can always fulfill the optimality condition, in practice there are some limitations due to the finite run of the correspondent axis.

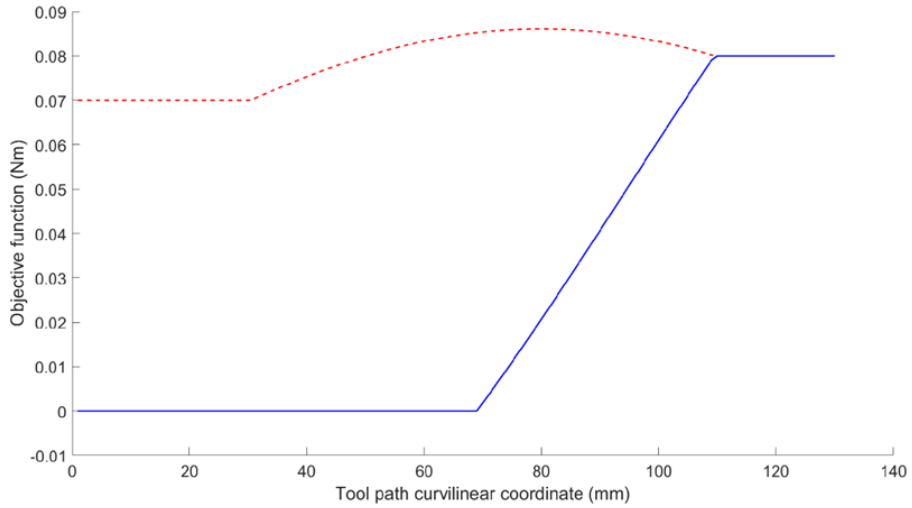


Figure 4.20: evolution of U_2 with respect to the curvilinear coordinate of the tool path

4.4.5 Evaluation of the proposed solution

Claims of the patent	Modified machine features	Potential advantages
Introduction of a redundant degree of freedom (actuated) in the kinematic chain	Improvement of the work space of the machine with same size	Wider number of workpiece typologies (bigger components)
		Decrease of the size of the machine
	Improvement of the kinematic-static performances of the machine	Mitigation of the dimensional machining errors
		Smaller amount of power required by motors
Introduction of an optimization procedure in the control logic of the machine, due to the introduction of a redundant link, (machine tools), or taking advantage of a redundant DoF already existing (robotic manipulators)	Improvement of the dynamic performances of the machine	Limitation of the inertia forces
		Reduction of machining times

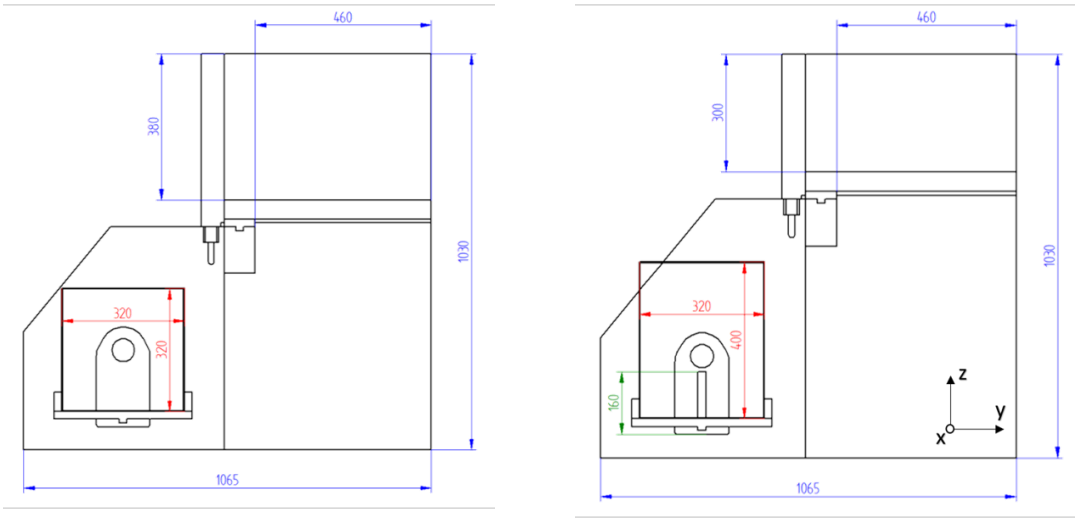
Table 4.3: potential advantages of implementation of the redundancy paradigm

Table 4.3 shows a resume of the potential advantages of the implementation of the redundancy paradigm. More specifically, all the potential advantages will now be discussed with reference to the confront between the RRTTT 5 axis milling machine, and the TRRTTT 6 axis milling machine.

Increase of the dimensions of workpieces and limitation of the size of the milling machine

Figure 4.30 shows the front view of two different layouts for a milling machine. On the left side, a standard RRTTT 5 axis milling machine is depicted, and on the right, it is shown the TRRTTT 6 axis redundant milling machine. The two machines have a similar kinematic chain, despite the fact that the latter integrates the redundant degree of freedom on the workpiece table.

The sizes of the two machines are the same, but, as it is shown in figure, the sizes of the workpieces are different. In the first case, the high of the maximum workable piece is 320 mm, while in the second case, the same dimension is 400mmm. Furthermore, even if the redundant mechanism has a further kinematic couple, the required stroke for the z axis actuator is smaller compared to the non-redundant mechanism. This is the effect of the enlargement of the workspace of the mechanism, due to the introduction of the redundant kinematic couple, integrated on the platform along the direction of the z axis.



(a) *(b)*

Figure 4.21: workspace increase due to the introduction of the redundant linear axis

Decrease of the positioning errors and limitation of the motor torques

The advantages in terms of mitigation of volumetric errors and containment of motor couples have already been largely illustrated in the previous subsection. Anyway, it is interesting highlight here that these two tasks are not in contradiction. On the contrary, it is possible to demonstrate that the condition of mitigation of volumetric error, is equivalent to an average optimality condition for the minimization of the actuator couple calculated on the angular range of all the possible directions for a force applied by the tool on a point of the workpiece. The brief demonstration of the previous statement may be found in appendix B.

Limitation of the inertia forces and decrease of production time

The introduction of the redundant degree of freedom can potentially allow the mitigation of another unwanted phenomenon. In figure 4.22 it is reported the classical situation in which, in order to impose a certain pose of the tool with to respect a fixed point of the work piece, it is necessary to provide a movement of the tool with respect the principal reference system of the machine tool.

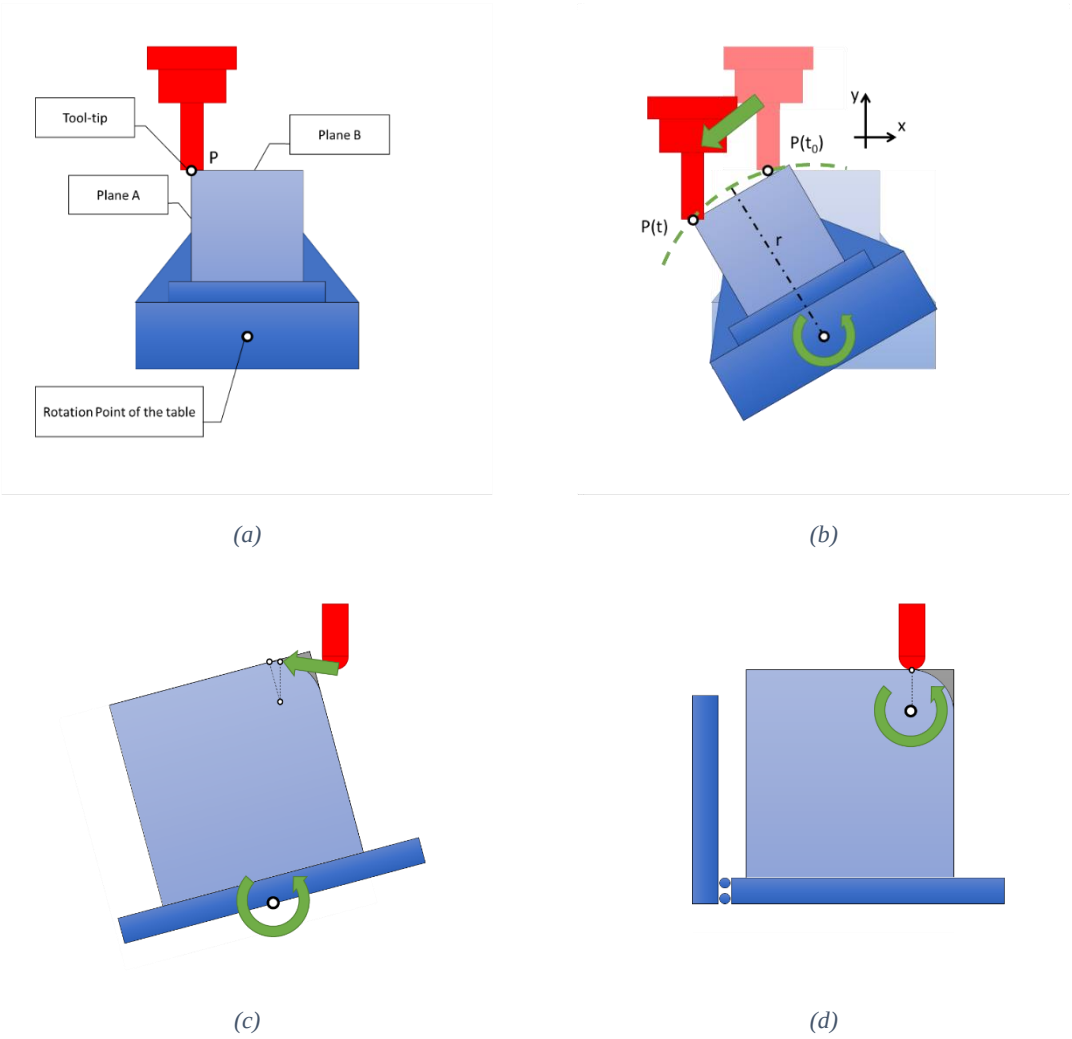
Referring to figure 4.22(a), the tool and the workpiece are at an initial mutual position and pose: in 5-axis milling machines, this configuration cannot vary, because it is direct consequence of the imposition of the better cutting conditions. Once the milling process start, the tool changes its relative position and pose with to respect the workpiece, and, as it is depicted in figure 4.22(b), depending to the position of the centre of rotation of the elements imposing the slope of the table, it is necessary to impose a certain tracking velocity to the tool, in order to allow its correct position with to respect the workpiece.

Furthermore, figure 4.23 represent the path of the tooltip during the following of the workpiece and the relative kinematic quantities. Let's suppose that the path along a quarter of circumference is covered by the tooltip with a constant velocity w. Associated to such motion can be defined an angular velocity $\dot{\theta}$, first time derivative of the angle θ:

$$\dot{\theta} = \frac{d\theta}{dt} = \frac{w}{r} \Rightarrow \theta = \frac{d\theta}{dt} \cdot t = \frac{w}{r} \cdot t \tag{4.48}$$

It is possible to express the horizontal component of the position of the tool as:

$$x = -r\cos\theta = -r\cos\left(\frac{w}{r}t\right) \tag{4.49}$$



(a) *(b)*

(c) *(d)*

Figure 4.22: representation of the motion of the tool due to the tracking of the workpiece (a), (b) and (c) and effect of the modify of the position of the turntable centre of rotation (d)

Furthermore, it is possible to define the other kinematic quantities such as the horizontal component of the velocity:

$$v_x = \frac{dx}{dt} = w sin\left(\frac{w}{r}t\right) = r\left(\frac{w}{r}\right) sin\left(\frac{w}{r}t\right) \tag{4.50}$$

the horizontal component of the acceleration:

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2} = \frac{w^2}{r} cos\left(\frac{w}{r}t\right) = r\left(\frac{w}{r}\right)^2 cos\left(\frac{w}{r}t\right) \tag{4.51}$$

and finally, the horizontal component of the jerk:

$$j_x = \frac{da_x}{dt} = \frac{d^3x}{dt^3} = -\frac{w^3}{r^2} sen\left(\frac{w}{r}t\right) = r\left(\frac{w}{r}\right)^3 sin\left(\frac{w}{r}t\right) \tag{4.52}$$

As it is well known by most of the operators, the non-null accelerations and jerks produce dynamical effects which affect the quality of the machining process.

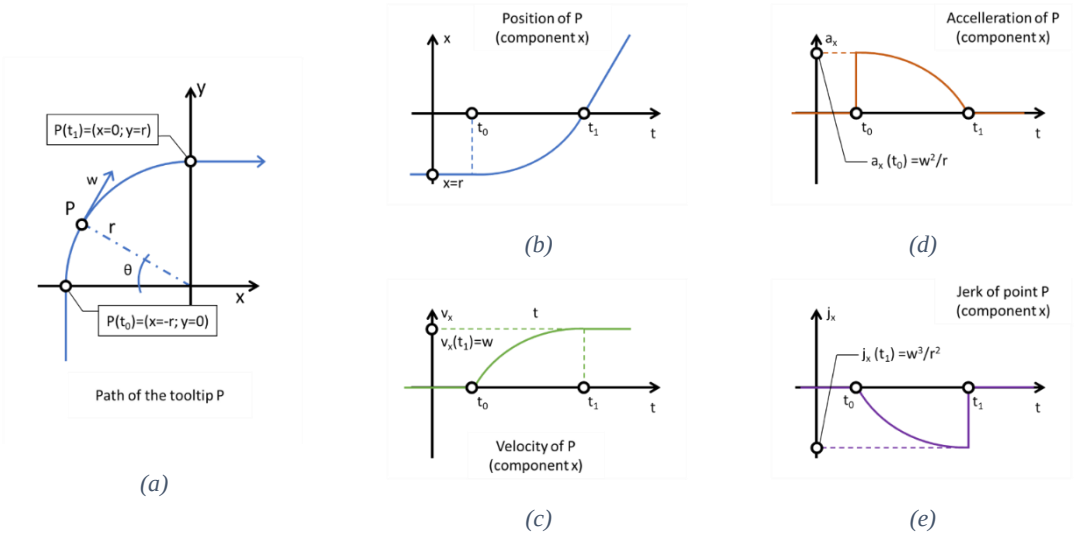


Figure 4.23: derivation of the kinematic quantities due to the tool tracking the workpiece

Equations (4.51) and (4.52) show that in order to limit the values of acceleration or jerk, it is necessary to minimize the distance r. Again, it is necessary to solve the optimization problem (4.22), obtaining two advantages: the mitigation of the position errors due to the angular pose of the platform, and a decrease of the dynamic requirements for the mechanism.

Difficulties in implementing redundancy paradigm

As a counterpart of the advantages described in the previous subsections, the introduction of a further degree of freedom in the kinematic chain of a mechanism implies some drawbacks as well.

The first difficulty regards the mechanical design. In fact, machine tools, such high productive 5 axis milling machines, are required to fulfil their tasks within a very strict tolerance field; for this reason, their mechanical structure is very stiff, and this means that the introduction of redundant degree of freedom should not affect this characteristic. Consequently, the enrichment of the kinematic chain implies a not trivial activity of mechanical design and a consequent rise of the of the production cost.

Another issue is the management of the control of the modified machine. As a matter of the fact, control systems for machine tools follow well established standards, and the modify of the mechanical structure implies a revamping of these subsystems.

Finally, it may be recalled that the Computer Aided Manufacturing (CAM) applications, which are the informatic tools for the generation of the motion laws of the of the machine tool, are standard software as well, and are not designed to manage more than 5 axis.

The result of a preliminary investigation of the issues reported above is that, even if there is the actual possibility of implementing redundancy in machine tools at a technical level, it is necessary to carry out cost/advantages analysis in order to evaluate the economic feasibility of the project.

4.5 Conclusions and future developments

This chapter was devoted to the description of the implementation of the redundancy paradigm in the machine design process. The study of productive means such as CNC milling machines and 3D printers has been identified as a field of application: the analysis of the state of the art revealed the importance in mitigating the volumetric errors due to the propagation of uncertainties in the angular positioning of the turntable motors.

The proposed solution is the result of the design methodology which provides for the implementation of an optimization procedure applied to a redundant kinematic structure. Actually, this approach got inspired by the two pillars described at the beginning of the present work: the biomimetic inspiration, which in this case is the conceptual idea of redundancy, and the definition of a constrained optimization problem.

It is well known by the theory of mechanisms that the introduction of additional degrees of freedom in the kinematic chain of a serial manipulator allows the fulfilment of different task requirements. Adding a redundant kinematic couple, the architecture of an existing machine causes the mathematical indeterminacy of the solution of the inverse kinematic problem: this means that, if the placement of the end effector of a manipulator according to a prescribed position and pose is the primary task, and if the mechanism is redundant, at least a second task can be fulfilled. The secondary task can be expressed by a wide variety of equations, which may be adopted as objective functions of a constrained optimization problem. The solution of the optimization problem leads to an improvement of specific indexes which characterize the performance of the mechanism.

Actually, the described process involves two stages of the design of the mechanism: the first stage regards the morphology of the structure of the machine; as a second stage, the control of the machine must be designed in order to provide a solution to the optimization process. This kind of design strategy, involving both design of structure and control is called integrated design, and a practical example has been provided by the case study of the TRRTTT 6 axis redundant milling machine. The performances of the proposed mechanism have been compared to the ones of the classical RRTTT 5 axis milling machine, by the means of a virtual models and all the results confirmed the effectiveness of the method.The future developments of this particular methodology are the definition of redundant variants of many other mechanisms, and the investigation of a larger number of objective functions in order to define and improve different performance indexes.

4.6 Bibliography

[57] Sciavicco, L., & Siciliano, B. (2012). Modelling and control of robot manipulators. Springer Science & Business Media.

[58] Bryan, J. B. (1979). The Abbe principle revisited: an updated interpretation. Precision Engineering, 1(3), 129-132.

[59] Patent Application WO/2019/243986. Available online at WIPO site https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2019243986.

[60] Caputi, A., & Russo, D. (2020). The optimization of the control logic of a redundant six axis milling machine. Journal of Intelligent Manufacturing, 1-13.

[61] Kiridena, V. F. P. M., & Ferreira, P. M. (1993). Mapping the effects of positioning errors on the volumetric accuracy of five-axis CNC machine tools. International Journal of Machine Tools and Manufacture, 33(3), 417-437.

[62] Cheng, Q., Zhao, H., Zhao, Y., Sun, B., & Gu, P. (2018). Machining accuracy reliability analysis of multi-axis machine tool based on Monte Carlo simulation. Journal of Intelligent Manufacturing, 29(1), 191-209.

[63] Lamikiz, A., De Lacalle, L. L., Ocerin, O., Díez, D., & Maidagan, E. (2008). The Denavit and Hartenberg approach applied to evaluate the consequences in the tool tip position of geometrical errors in five-axis milling centres. The International Journal of Advanced Manufacturing Technology, 37(1-2), 122-139.

[64] Ziegert, J. C., & Kalle, P. (1994). Error compensation in machine tools: a neural network approach. Journal of Intelligent Manufacturing, 5(3), 143-151.

[65] Suh, S. H., Lee, E. S., & Jung, S. Y. (1998). Error modelling and measurement for the rotary table of five-axis machine tools. The International Journal of Advanced Manufacturing Technology, 14(9), 656-663.

[66] Cheng, Q., Feng, Q., Liu, Z., Gu, P., & Zhang, G. (2016). Sensitivity analysis of machining accuracy of multi-axis machine tool based on POE screw theory and Morris method. The International Journal of Advanced Manufacturing Technology, 84(9-12), 2301-2318.

[67] Khan, A. W., & Wuyi, C. (2010). Systematic geometric error modeling for workspace volumetric calibration of a 5-axis turbine blade grinding machine. Chinese Journal of Aeronautics, 23(5), 604-615.

[68] Fu, G., Fu, J., Xu, Y., Chen, Z., & Lai, J. (2015). Accuracy enhancement of five-axis machine tool based on differential motion matrix: geometric error modeling, identification and compensation. International Journal of Machine Tools and Manufacture, 89, 170-181.

[69] He, K., Zhang, Q., & Hong, Y. (2019). Profile monitoring based quality control method for fused deposition modeling process. Journal of Intelligent Manufacturing, 30(2), 947-958.

[70] Stryczek, R. (2016). A metaheuristic for fast machining error compensation. Journal of Intelligent Manufacturing, 27(6), 1209-1220.

[71] Kreng, V. B., Liu, C. R., & Chu, C. N. (1994). A kinematic model for machine tool accuracy characterisation. The International Journal of Advanced Manufacturing Technology, 9(2), 79-86.

[72] Kiridena, V. F. P. M., & Ferreira, P. M. (1993). Mapping the effects of positioning errors on the volumetric accuracy of five-axis CNC machine tools. International Journal of Machine Tools and Manufacture, 33(3), 417–437.

[73] Ji,W., &Wang, L. (2019). Industrial robotic machining: A review. The International Journal of Advanced Manufacturing Technology, 103(1–4), 1239–1255.

[74] Tao, J., Qin, C., Xiao, D., Shi, H., Ling, X., Li, B., & Liu, C. (2020). Timely chatter identification for robotic drilling using a local maximum synchrosqueezing-basedmethod. Journal of Intelligent Manufacturing, 31(5), 1243–1255.

[75] Xiong, G., Ding, Y., & Zhu, L. (2019). Stiffness-based pose optimization of an industrial robot for five-axis milling. Robotics and Computer-Integrated Manufacturing, 55, 19–28.

[76] Xiao, W., & Huan, J. (2012). Redundancy and optimization of a 6R robot for five-axis milling applications: Singularity, joint limits and collision. Production Engineering, 6(3), 287–296.

[77] Chen, G., Liang, Y., Sun, Y., Chen, W., & Wang, B. (2013). Volumetric error modeling and sensitivity analysis for designing a five-axis ultra-precision machine tool. The International Journal of Advanced Manufacturing Technology, 68(9–12), 2525–2534.

[78] Pedersen, D. B., Eiríksson, E. R., Hansen, H. N., & Nielsen, J. S. (2016). A self-calibrating robot based upon a virtual machine model of parallel kinematics. Virtual and Physical Prototyping, 11(3), 227–234.…

[79] Calignano, F. (2018). Investigation of the accuracy and roughness in the laser powder bed fusion process.Virtual and Physical Prototyping, 13(2), 97–104.

[80] He, K., Zhang, Q.,&Hong, Y. (2019). Profilemonitoring based quality control method for fused deposition modeling process. Journal of Intelligent Manufacturing, 30(2), 947–958.

[81] Song, X., Pan, Y., & Chen, Y. (2015). Development of a low-cost parallel kinematic machine for multidirectional additive manufacturing. Journal of Manufacturing Science and Engineering, 137(2), 021005.

[82] Flynn, J. M., Shokrani, A.,Newman, S. T.,&Dhokia,V. (2016).Hybrid additive and subtractive machine tools—Research and industrial developments. International Journal of Machine Tools and Manufacture, 101, 79–101.

5 An ontological tool for the analysis and synthesis of compliant mechanisms

5.1 Definition of compliant mechanism and scope of the chapter

By definition, a compliant mechanism is a monolithic structure whose function is transmitting mechanical power from a point (or different points), to another (or others), by the mean of the deformation of its parts. Compliant mechanisms own features of both mechanisms and structures; as it had been presented in chapter 4, mechanisms are usually schematized as a set of perfectly rigid members, linked together by kinematic couples; the model and the control of a mechanisms is based on creating a relation between a fixed reference system, and a reference system bound to an end effector. On the other hand, according to the theory reported in chapter 3, structures are continuous systems, which occupy different regions of Euclidean space, depending on the forces that are applied to it; when a force system is applied to a deformable body, the unloaded configuration and the loaded configuration may be put in relation referring to a displacement field.

The scope of this chapter is the creation of a common model including both structures and mechanisms features, and the construction of an ontological framework, in order to provide an actual tool for the analysis and the design of the compliant mechanisms.

5.1.1 Continuum model vs. discrete model

In literature, it is possible to find two main methodologies for the synthesis of compliant mechanisms [83]. The first one is the so called “pseudo rigid body” model, and it has been disclosed by Howell [84]: he proposed to extend the theory for the analysis and synthesis of the kinematics of the rigid body to compliant mechanisms, by the introduction of well characterized flexible elements. The result is a concentrated parameters model which, on the one hand, has a well-established theoretical background, and is easy to implement, and on the other hand, provides a good approximation of the behaviour of the actual deformable structure.

The second methodological framework is the continuum structures topology optimization [85]. Here the idea, as explained in chapter 3, is to reformulate the synthesis problem in the problem of the identification of the ideal material distribution by the means of the solution of a constrained optimization problem, based in a finite element discretization of the design space. In this case, differently by the minimum compliance problem, the objective function will be related to displacements and forces applied to different points of the continuum.

The analysis of the later method shall be the subject of the first part of the chapter. The former will be discussed in the second part, and it will be used in the implementation of the ontological framework, constituting the main outcome of this part of the research.

5.2 Synthesis of continuum compliant mechanisms

It may be recalled from chapter 3 that, in order to setup the topology optimization problem for continuum structures, it is necessary to operate some choices:

1) as a first stage, it is necessary to define a design domain; as an example, ground structures can be taken in account, in order to generate truss assembly; alternatively, it is possible to discretize a certain domain space due to a finite element schema;
2) the second step is the definition of the material characterization, such as binary or SIMP interpolation schema, introduced in chapter 3;
3) furthermore, the optimization problem must be set up by the specification of the objective function, representing the quantity to minimize or maximize, such in the case of the minimization of the compliance for the structures.

Actually, a synthesis procedure for the compliant mechanisms requires all these features, even if the definition of a, let us say, functional requirement for the fulfilment of point (3), is a more complex task for compliant mechanisms. In fact, it is further required:

- to define a mechanism model; this can be done by a functional definition of the task of the mechanism. As it will be better described later on, it is possible to require the maximization of the displacement of a certain point, or, as an alternative, the force transferred by the input port to output port. Depending on the specific formulation, the resulting topology may vary significantly;

- once the mechanism model is defined, it is necessary to define an objective function that will be the mathematical counterpart of the mechanism model: similarly to topology optimization for structural purposes (i.e. choice of minimal compliance or resistance criterion), this step may deeply affect the result of the optimization process.

The introduction of these further steps depends by the fact that the definition of compliant mechanism is not a trivial task. The reason is that it involves two different requirements in contrast one another: the first requirement consists in providing enough flexibility to allow the effective displacement of the output port (end effector). The second requirement is that the monolithic structure of the compliant mechanism must be stiff enough in order to allow the transmission of the force applied in correspondence of the input port, to the output port.

5.2.1 Stiffness vs. flexibility: strain energy and mutual strain energy

Previously, as it can be suggested by intuition, it had been stated that compliant mechanisms are particular devices which must fulfil two different and conflicting requirements: the maximization of the displacement of the relative displacement of some parts, in order to fulfil the kinematic requirements, and the maximization of the stiffness of the structure, in order to allow the efficient transfer of the forces. Consequently, it is necessary to define the mathematical quantities related to the concepts of stiffness and flexibility, feasible to be implemented in an objective function for the optimization algorithm.

Strain energy

As it had been already discussed in chapter 3, one of the principal indexes characterizing the behaviour of a structure is the mean compliance, or, in other words, the work done by the external forces according to the equation 3.22, it represents the deformation of the structure itself:

$$\int_{\Omega_E} \left(\delta \underline{\varepsilon}\right)^T \cdot \underline{\sigma}\, d\Omega = \int_{\Omega_E} \left(\delta \underline{v}\right)^T \cdot \underline{b}\, d\Omega + \int_{\Gamma} \left(\delta \underline{v}\right)^T \cdot \underline{t}\, d\Gamma + \sum_{i=1}^{4} \delta \underline{v}\big|_{\underline{x}=\underline{x}_i} \cdot f_i$$

As it is clearly stated by Howell [84], the mean compliance is directly related to the stiffness of the structure: the lower is the former, the higher is the latter. Moreover, mean

compliance expresses the total amount of elastic energy stored in the structure, because it is the double of the total strain energy of the system. In other words, if only a force is applied to the structure, it is possible to write:

$$\int_{\Omega_E} \underline{\varepsilon}^T \cdot \underline{\sigma}\, d\Omega = f_{in} \cdot \delta_{in} = MC = 2 \cdot SE \tag{5.1}$$

being MC the mean compliance of the structure due to the force applied, being the double of the strain energy

$$SE = \int_{\Omega_E} \frac{1}{2} \cdot \underline{\varepsilon}^T \cdot \underline{\sigma}\, d\Omega \tag{5.2}$$

Actually, the mean compliance is the first parameter considered for the characterization of the compliant mechanism. Adopting the finite element framework for the discretization of the continuum, it is possible to calculate the mean compliance MC if the stiffness matrix is known by determining the nodal displacements:

$$\underline{\underline{K}} \cdot \underline{U} = \underline{F_{in}} \tag{5.3}$$

where the vector $\underline{F_{in}}$ is a vector identically null except for the component corresponding to the force f_{in}; once the nodal displacements are known, MC can be easily calculated:

$$MC = \underline{F_{in}} \cdot \underline{U} = \underline{U}^T \cdot \underline{\underline{K}} \cdot \underline{U} = f_{in} \cdot \delta_{in} \tag{5.4}$$

Mutual strain energy

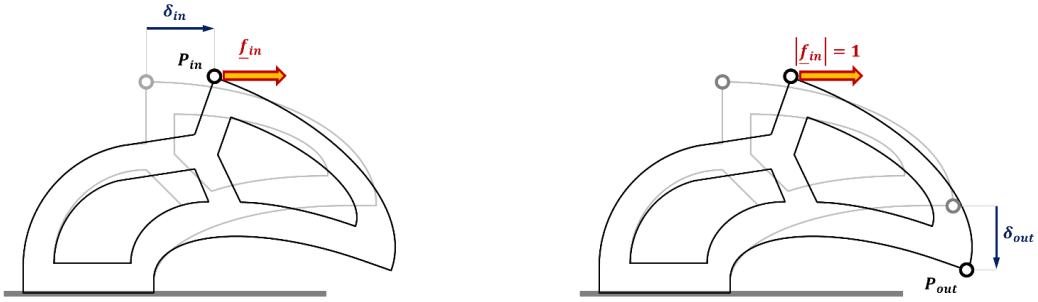
Actually, if a structure is subject to a single load, mean compliance or total elastic strain energy allow to directly calculate the displacement of the point of application of the force along its direction, and this is the measure of how much is stiff the structure in this particular load condition. The application point of the force is usually called input port.

On the other hand, in the synthesis of compliant mechanism, the main goal is allowing the possibility of realizing a large displacement of a second point, which is actually the functional requirement of the mechanism. Such displacement can be easily found by the application of the principle of the virtual works. Let us assume to have two different systems: a first system, usually called actual system, is the one described above, where the structure is subject to a force $\underline{f_{in}}$ applied in the input port P_{in}. On the other hand, a second system is defined, called virtual system, where the structure is subject to a dummy

load $\underline{f}_{out} \equiv \underline{f}_{dummy}$, being $\left|\underline{f}_{dummy}\right| = 1$, in the application point P_{out}. If we indicate the tension field generated by the dummy load with the symbol $\underline{\sigma}_{dummy}$, then, due to the principle of the virtual works, it is possible to write the following integral expression:

$$\int \underline{\varepsilon}^T \cdot \underline{\sigma}_{dummy} = f_{out} \cdot \delta_{out} = \delta_{out} = MSE \tag{5.5}$$

and this is the general expression of the so-called mutual strain energy MSE.



(a) *(b)*

Figure 5.1: displacement at input port (a) and output port (b), due to a dummy load applied at input port

Adopting the finite element framework for the discretization of the continuum, it is possible to calculate the nodal displacements for the structure in the two cases, when the structure is subject to the actual load, and when it is subject to the dummy load:

$$\begin{aligned} \underline{\underline{K}} \cdot \underline{U} &= \underline{F}_{in} \\ \underline{\underline{K}} \cdot \underline{V} &= \underline{F}_{out} \end{aligned} \tag{5.6}$$

recalling that the vector $\underline{F}_1$ is a vector identical equal to zero, except the component corresponding to the degree of freedom corresponding to the input force, which is equal to 1, and the vector $\underline{U}$ is correspondent vector of the nodal displacements.

Furthermore, let us consider a second load case, where the unit (dummy) load $\underline{F}_2$ is applied to the same structure at output port; $\underline{V}$ is the nodal displacements vector, according to:

$$\underline{\underline{K}} \cdot \underline{V} = \underline{F}_2 \tag{5.7}$$

Applying the principle of the virtual works it is possible to write:

$$\underline{F_2} \cdot \underline{U} = \underline{V}^T \cdot \underline{\underline{K}} \cdot \underline{U} = 1 \cdot \delta_{out} = \delta_{out} \tag{5.8}$$

where the displacements are the ones corresponding to the first load case, and the forces (internal forces and dummy load) are the ones corresponding to the second load case. As it is defined, the mutual strain energy represents an index which describes the capability of the structure to allow the displacement along that particular degree of freedom when subject to a prescribed load.

Geometrical and mechanical advantage

Strain energy and mutual strain energy are the quantities related to the contrasting requirements stiffness and flexibility of the structure. It is possible to define other three quantities, related, this time, to the kinematic-static duality, introduced in the chapter 4. In fact, it is possible to indicate the efficiency of a mechanism in transferring mechanical power from a point (input port), to another point (output port) by the measurement of the geometrical advantage (GE), the mechanical advantage (MA), and the work ratio (WR) [86].

As a definition, geometric advantage is the ratio of the displacement of the output port and the displacement of the input port due to the forces acting on the structure. The mechanical advantage is defined as the ratio of the force applied to the input port, and the one applied to the output port. Finally, the work ratio is the product of the geometric advantage and mechanical advantage, or, equivalently, the ratio of the work done by the force applied on the output port during its displacement, and the work done by the force at the input port. As an important remark, let us notice that, differently by (ideal) mechanisms composed by rigid bodies linked by ideal joints, the energy provided at the input port is not equal to the energy available at the output port: in fact, part of the mechanical energy is converted in potential elastic energy. Furthermore, this condition can be expressed in terms of mechanical advantage as well [87].

5.2.2 Set up of the optimization problem: objective functions and task model

In order to set up a topology optimization method suitable for the synthesis of compliant mechanism, it is necessary to identify a convenient objective function, or, more precisely, a set of objective functions [88], able to describe the functional requirements. In fact, as

outlined previously, differently by the synthesis of structures, the design of the compliant mechanism is a trade-off between two functional objectives.

Objective function

As it had been stated in chapter 3, the topology optimization problem is re-formulated as the problem of finding the best $\underline{\underline{E}} \in \underline{\underline{E}}_{adm}$ tensor field, which characterizes the material distribution, and minimizes a certain functional, recalling that, in this general definition, $\underline{\underline{E}}$ is a field, so the quantity to minimize is defined by a functional. In order to adopt topology optimization for the synthesis of compliant mechanisms, it is necessary to formalize the relation between the elastic properties of the structure, and the kinematic requirements for the mechanism.

In the previous subsection, there have been introduced four indexes that may help us to describe the characteristics and the behaviour of the compliant mechanism. Strain energy, mutual strain energy, geometrical advantage and mechanical advantage are the main elements for the formulation if the objective function for the optimization problem.

As it stated by Howell [84], a first, natural formulation of an objective function is a linear combination of SE and MSE

$$-\alpha MSE + (1-\alpha)SE \tag{5.9}$$

or, as an alternative, the ratio of MSE and SE

$$\frac{-MSE}{SE} \tag{5.10}$$

At a first glance, these may seem a convenient objective function for a synthesis process, but, unfortunately, this problem setup doesn't lead to a feasible solution. In fact, if no force is applied to the output port, the topology which maximize the displacement of the output, or the geometrical advantage, ratio of output port displacement and input port displacement, is a structure in which there is not a physical connection between the two ports. This can be explained intuitively by simply considering that, if an algorithm is required to maximize the displacement of the output port, which is actually equivalent to MSE, and that maximize SE at the same time, will naturally generate a topology in which the input port and the output port are physically disconnected. An example of this resulting topology is depicted in figure 5.2. Moreover, the displacement of the output port will be maximized by the creation of a structure having stiffness equal to zero. The

theoretical background which is at the base of this consideration will be shown in the next subsection, when kinetic/elastic model will be introduced.

A commonly used solution for this drawback is the modification of the structure mode introducing a linear spring, disposed along the degree of freedom of interest of the output port. As it had been largely showed in literature, this allows the optimization algorithm to obtain a consistent and connected structure. This should not surprise: in fact, it is reasonable to think that, in order to identify the ideal topology for the mechanism/structure, it is important specify the way it interact with the environment, defining the nature of the forces acting on its input and output ports.

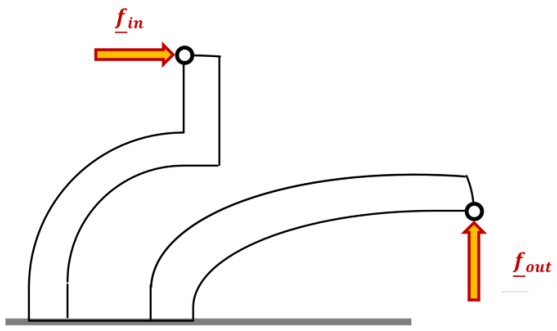


Figure 5.2: solution of continuum compliant mechanism synthesis characterized by the disconnection between the input port and the output port

Task model

As it addressed by Sighmund [85] is possible to distinguish different modes of interaction of the structure with the environment. Here are some of the possible cases, as reported in figure 5.3:

a) no resistance force: the structure is only subject to the load applied at the input port; in this case, possible objective functions are the displacement of the output port, or the geometrical advantage, since the mechanical advantage may not be calculated, because there is not the transfer of mechanical energy, and all the work done by the force is all stored as elastic energy due to the deformation of the structure. As it has been already highlighted, will be better shown later, the use of this model can't lead to a reasonable solution of the optimization process;

b) the compliant mechanism applies a force on a much stiffer body: in this case, the displacement of the output port is null, and consequently the geometric advantage is null as well; on the other hand, a reaction force is applied by the body to the compliant mechanism, and this means that the mechanical advantage can be taken in account as objective function of the optimization process;

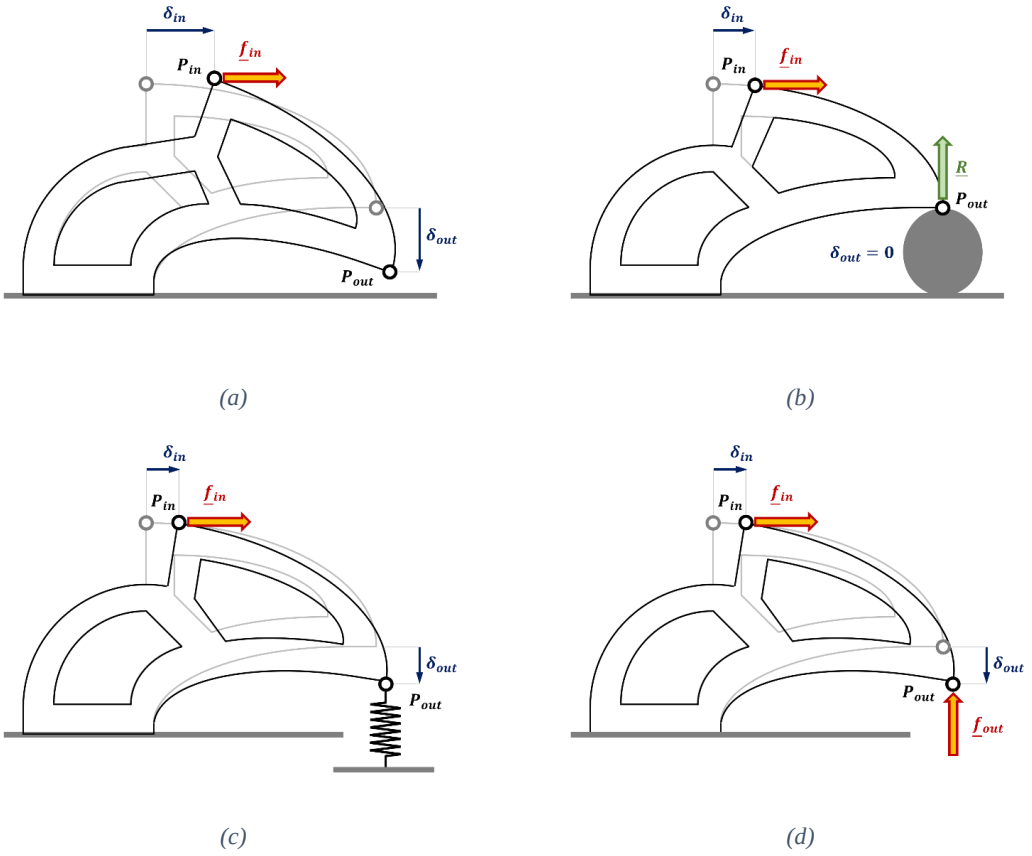


Figure 5.3: examples of interaction modes between the compliant mechanism and the environment

c) the third case is the so-called spring model: as it will be shown later, this problem setup allows the convergence of the optimization problem to a connected topology, allowing the compliant mechanism to actually transfer mechanical power from the input port to the output port. Anyway, as an anticipation of what will be better explained later on, the adoption of this task model leads to a structure showing the so-called actual hinges issue: this means that the algorithm leads to a solution in which all the flexibility of the structure/mechanism is concentrated in precise points, connecting other parts characterized by a higher stiffness. This outcome can be

explained by the fact that, actually, this kind of solution is the natural way to minimize the compliance of the structure, and maximize the displacement of the output port at the same time;

d) the fourth case considers the two forces applied to the input and output port. As it will be shown later, this model allows to change the approach in defining the objective function, which can be defined in terms of stiffness characteristics of the structure.

As a remark, it can be said that the enlisted models are only a selection of the ones which can be found in literature: some researchers proposed the application of a linear spring on bot input and output port; another example considers a gap between the undeformed position of the output port, and the actual point of interaction between the mechanism and the environment.

5.2.3 Stiffness model

As it had been anticipated previously, if the free output port is adopted, simply maximizing output displacement or geometric advantage leads to meaningless topologies: in order to formally demonstrate this statement, it is necessary to introduce the idea of the stiffness model of the compliant mechanism.

Again, let us assume to discretize the continuum using the finite elements framework, and suppose to define two distinct systems, as shown in figure 5.4. As it had been done in the previous section, the two systems are substantially the same structure, subject to two different forces. Even if in literature the two systems are called the "real" system, and the "virtual" system, here it will be used the more general notation system i, being i=1,2. It is assumed that, for the system i, the structure is subject to the force $\underline{f_i}$ applied in the point P_i; due to the effect of the load, the deformation of the structure is represented by the vector of the nodal displacements $\underline{V_i}$, and in particular δ_{ij} is the displacement of the point P_i due to the force $\underline{f_j}$ (along the direction of the force $\underline{f_j}$). As a further hypothesis, the force applied is a dummy load, so that $\left|\underline{f_i}\right| = 1$.

The compliance of the system 1 is equivalent to the strain energy stored in the structure, and it can be written as:

$$MC_1 = f_1 \cdot \delta_{11} = \delta_{11} = \underline{V_1}^T \cdot \underline{\underline{K}} \cdot \underline{V_1} \tag{5.11}$$

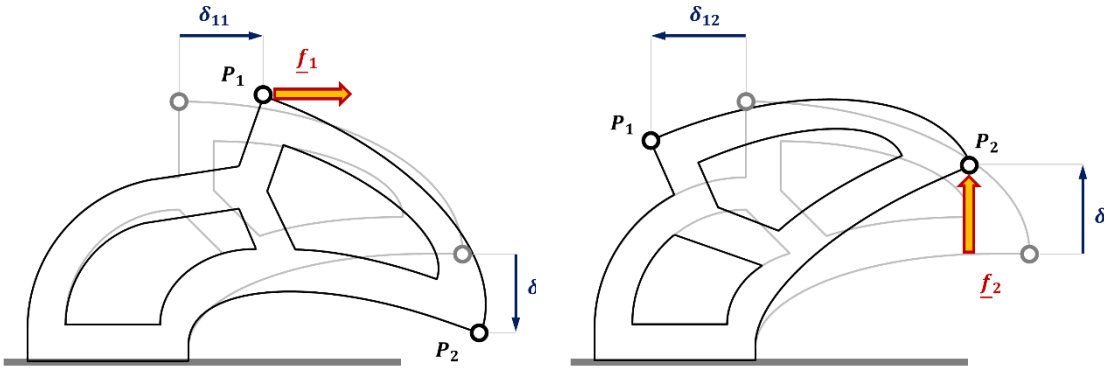
and, similarly, the compliance of the system 2 is:

$$MC_2 = f_2 \cdot \delta_{22} = \delta_{22} = \underline{V_2}^T \cdot \underline{\underline{K}} \cdot \underline{V_2} \tag{5.12}$$

$\underline{\underline{K}}$ is the rigidity matrix of the structure, and recalling that, by definition, the rigidity matrix of the structure must satisfy the following equations

$$\underline{\underline{K}} \cdot \underline{V_i} = \underline{F_i} \quad ; \quad i = 1,2 \tag{5.13}$$

where $\underline{F_1}$ and $\underline{F_2}$ are two vectors identically null except the component of the degree of freedom corresponding respectively to the forces $\underline{f_1}$ and $\underline{f_2}$ not null components: let us say that these are the k-th and m-th component respectively.



(a) *(b)*

Figure 5.4: displacement of the input port due to the input dummy load (a), and output dummy load (b)

Furthermore, it is possible to write the mutual strain energy of the system i, due to a dummy load applied in the port P_j (being $i \neq j$): recalling that $\left|\underline{f_1}\right| = \left|\underline{f_2}\right| = 1$, due to the Maxwell's theorem it follows that:

$$MSE = MSE_1 = \underline{V_1}^T \cdot \underline{\underline{K}} \cdot \underline{V_2} = \delta_{12} = MSE_2 = \underline{V_2}^T \cdot \underline{\underline{K}} \cdot \underline{V_1} = \delta_{21} \tag{5.14}$$

According to what has been introduced by Wang [89] without any consequence on the generality of the reasoning, it is possible to rearrange the rows and the columns of the matrix $\underline{\underline{K}}$, so that the first component of the generic nodal displacements vector

corresponds to the non-null component of the force $\underline{f_1}$ (originally the k-th component), and the second component to the non-null component of the force $\underline{f_2}$ (originally the m-th), so that:

$$\frac{1}{MC_1 \cdot MC_2 - MSE^2} \cdot \begin{vmatrix} MC_2 & -MSE \\ -MSE & MC_1 \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix} \tag{5.15}$$

for a generic structure subject to two single components forces. The demonstration of the equation 5.15 can be found in the appendix C.

5.2.4 Example of optimization setup and adoption of the spring model

Let us consider the optimization problem in which the objective function is the geometric function, according to the task model the task depicted in figure 5.3(a). As it had been shown by Wang [89], both maximization of the geometric or mechanic advantage, lead to a disconnected topology. In fact, the definition of geometric advantage is the following

$$GE = \left|\frac{u_2}{u_1}\right| = \left|\frac{u_{out}}{u_{in}}\right| \tag{5.16}$$

and, in the case of free output port task model, the displacements of the input and output ports depend only on the force applied at the input port:

$$\begin{cases} f_1 \neq 0 \\ f_2 = 0 \end{cases} \quad \Rightarrow \quad \begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} f_{in} \cdot MC_1 \\ f_{in} \cdot MSE \end{vmatrix} = \begin{vmatrix} f_{in} \\ 0 \end{vmatrix} \tag{5.17}$$

Consequently, the geometric advantage may be written as follows

$$\begin{aligned} GE = \left|\frac{u_{out}}{u_{in}}\right| = \left|\frac{MSE}{MC_1}\right| &= \left|\frac{\frac{MSE}{MC_1 \cdot MC_2 - MSE^2}}{\frac{MC_1}{MC_1 \cdot MC_2 - MSE^2}}\right| \\ &= \left|\frac{H_s}{H_2}\right| \quad if \quad MC_1 \cdot MC_2 - MSE^2 \neq 0 \end{aligned} \tag{5.18}$$

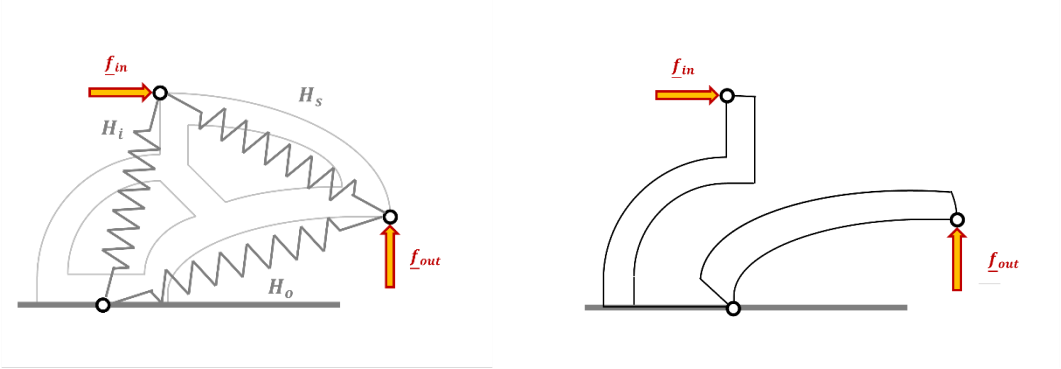
This means that the trivial solution for the maximization of GE corresponds to the topology such that $MC_1 \to 0$. Consequently, the matrix $\underline{\underline{R}}$ is such that

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \quad \to \quad \begin{vmatrix} H_1 & H_s \\ H_s & 0 \end{vmatrix} \quad \Longrightarrow \quad \begin{vmatrix} H_1 & H_s \\ H_s & 0 \end{vmatrix} \cdot \begin{vmatrix} u_{in} \\ u_{out} \end{vmatrix} = \begin{vmatrix} f_{in} \\ 0 \end{vmatrix} \tag{5.19}$$

In order to have non null input displacements $u_{in} \neq 0$, from second equation leads to having a singular matrix:

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \rightarrow \begin{vmatrix} H_1 & H_s \\ H_s & 0 \end{vmatrix} \rightarrow \begin{vmatrix} H_1 & 0 \\ 0 & 0 \end{vmatrix} \tag{5.20}$$

The physical meaning of the singularity of the matrix $\underline{\underline{R}}$ is depicted in figure 5.5: figure 5.5(a) shows the interpretation of the condensed parameters R_1, R_2 and R_3, being R_1 the connection between the input port and the ground, R_2 the connection between the output port and the ground, and R_s the connection between the input port and the output port.



(a) *(b)*

Figure 5.5: compliant mechanism condensed parameters model (a), disconnected geometry for R_2=0 (b)

According to this schema, the condition $R_s = 0$ implies the disconnection between the input and output port, as reported in figure 5.2, and, furthermore, the condition $R_2 = 0$ a rigid body displacement mode, as shown in figure 5.5(b). As it had been largely reported in literature, spring task model has been successfully applied in order to avoid the disconnection between the input and output port. This is due by the fact that the global matrix $\underline{\underline{H}}$ is modified by the introduction of the external spring as follows

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 + k_0 \end{vmatrix} \cdot \begin{vmatrix} u_{in} \\ u_{out} \end{vmatrix} = \begin{vmatrix} f_{in} \\ 0 \end{vmatrix} \tag{5.21}$$

and this implies that

$$GE = \left| \frac{u_{out}}{u_{in}} \right| = \left| \frac{H_s}{H_2 + k_0} \right| \tag{5.22}$$

This means that the trivial solution for the maximization of GE corresponds to the topology such that $H_2 + k_0 \rightarrow k_0$, the matrix $\underline{\underline{R}}$ is such that

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 + k_0 \end{vmatrix} \rightarrow \begin{vmatrix} H_1 & H_s \\ H_s & k_0 \end{vmatrix} \Rightarrow \begin{vmatrix} H_1 & H_s \\ H_s & k_0 \end{vmatrix} \cdot \begin{vmatrix} u_{in} \\ u_{out} \end{vmatrix} = \begin{vmatrix} f_{in} \\ 0 \end{vmatrix} \tag{5.23}$$

This means that, in order to have not null input displacements $u_{in} \neq 0$, it is no more necessary that $H_s = 0$, the matrix $\underline{\underline{H}}$ is no more singular.

5.2.5 The actual hinges issue

The adoption of the spring model is an effective way to overcome the singularity if the matrix $\underline{\underline{H}}$, and consequently allows to obtain a connected topology for the compliant mechanism; nevertheless, the external spring task model is affected by another issue that may lead to a not desired result in terms of topology. In fact, it is possible to demonstrate that the optimization algorithms based on this approach are affected by the so-called actual hinges issue: it means that the resulting topology is composed by zones with high rigidity, very massive and stiff, connected by small zones having high flexibility.

This result is easily explained thinking about the fact that the optimization is carried out trying to realize a trade-off between two contrasting goals: on one hand, the stiffness of the structure should be maximized by the minimization of the compliance, and, on the other hand, the displacement of the output port is required to be as large as possible; thinking about the fact that compliance is not a distributed quantity, but is the sum of the compliance of all the elements, it is easy to understand that the optimization process naturally leads to a structure which is as close as possible to a rigid body mechanism, having links, let us say, infinitely rigid, connected by rotational kinematic couples.

Under a mathematical point of view, according to Wang [89], even if the adoption of the artificial introduction of an external spring leads the optimization process to create a structure having $H_s \neq 0$, the optimal solution anyway requires that $H_2 = 0$. This means that the global stiffness matrix of the structure has the determinant equal to zero:

$$det\left(\underline{\underline{K}}\right) = det \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \begin{vmatrix} \cdots & \underline{k}_{1s}{}^{T} & \cdots \\ \cdots & \underline{k}_{2s}{}^{T} & \cdots \end{vmatrix} \\ \begin{vmatrix} \vdots & \vdots \\ \underline{k}_{1s} & \underline{k}_{2s} \\ \vdots & \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \cdots \quad \underline{\underline{k}}_{ss} \quad \cdots \\ \vdots \end{vmatrix} \end{vmatrix} = 0 \tag{5.24}$$

It may be noticed that this computation is caried out without the external springs, which by the mechanism side is an external system applying a force. Recalling that the elastic energy stored in a structure is the quadratic form

$$SE = \begin{Vmatrix} u_1 \\ u_2 \\ \vdots \\ \underline{u}_s \\ \vdots \end{Vmatrix}^T \cdot \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \begin{vmatrix} \cdots & \underline{k}_{1s}{}^T & \cdots \\ \cdots & \underline{k}_{2s}{}^T & \cdots \end{vmatrix} \\ \begin{vmatrix} \vdots \\ \underline{k}_{1s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \underline{k}_{2s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \cdots \ \underline{\underline{k}}_{ss} \ \cdots \\ \vdots \end{vmatrix} \end{vmatrix} \cdot \begin{Vmatrix} u_1 \\ u_2 \\ \vdots \\ \underline{u}_s \\ \vdots \end{Vmatrix} \tag{5.25}$$

it can be said that, if the determinant of the matrix is null, consequently the stiffness matrix is semi-definite positive, and this implies that a displacement mode $\underline{\bar{U}} = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n)^T$ exists such that:

$$\overline{SE} = \underline{\bar{U}}^T \cdot \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \begin{vmatrix} \cdots & \underline{k}_{1s}{}^T & \cdots \\ \cdots & \underline{k}_{2s}{}^T & \cdots \end{vmatrix} \\ \begin{vmatrix} \vdots \\ \underline{k}_{1s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \underline{k}_{2s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \cdots \ \underline{\underline{k}}_{ss} \ \cdots \\ \vdots \end{vmatrix} \end{vmatrix} \cdot \underline{\bar{U}} = 0 \tag{5.26}$$

But, on the other hand, this particular mode, which corresponds to a null elastic potential energy stored, correspond to a de facto rigid body mechanism, which actually does not store any elastic energy, or, in other words, a structure provided of the facto hinges.

An interesting remark regards the fact that there is a certain number of researches [90] that investigate the possibility of adopting continuum optimization strategies based on the spring model as inspiration for the design of mechanisms with rigid links, or, alternatively, the design of compliant mechanisms with concentrated deformations. It can be said anyway that, in general, distributed deformation is a desirable feature in the design of compliant mechanism, mostly for issues relate to the resistance and fatigue.

5.2.6 Review of the most common objective functions for the synthesis of compliant mechanisms

According to the analyses caried out in the previous subsections, some of the main features that should be required to a continuum topology optimization method for the synthesis of compliant mechanisms have been identified:

- implement a task model, and an objective function, or a set of objective functions, which can mathematically represent the functional requirement;
- ensure a connected topology for the compliant mechanism able to transfer mechanical power from the input port to the output port;
- avoid the generation of actual hinges, or, equivalently, promote the generation of a compliant mechanism with distributed compliance.

Accomplish these conditions is not an easy task, for the issues affecting continuum topology optimization, which have been highlighted in chapter 3, and, let us say, the difficult definition of the functional tasks. For these reasons, many researchers investigated many different approaches: some of them have been already cited, like and Wang [83][89], Howell [84], and Sighmund [85] which investigated the implementation of the geometric and mechanical advantage, in the case of different task models, and in particular the spring model. Nevertheless, many others are the contributions to the field: Ansola et al. [91] proposed a BESO like approach adopting the mechanical advantage as objective function and the spring model; Lau et al. [86] have been already cited because they proposed, as objective function, the use of the work ratio $WR \equiv MA \cdot GA$ coupled with the spring model. On the other hand, Li [92] in his doctoral thesis, among others, reported the use of the work ratio, but adopting two springs, one connected to the input port, and the other connected to the output port. This double spring model has been used by Liu et al. [93], while the adopted objective function is the simple output port displacement. Furthermore, the ratio of the geometric advantage and the input strain energy $\frac{GA}{SE}$ was proposed again by Li et al. [94].

All these approaches were based on the combination of forces and displacements applied to the input and output ports, as presented previously, but, as highlighted by Huang, Li, Zhou and Xie [95], which disclosed the objective function

$$-u_{OUT} + \lambda(C^* + C) \tag{5.27}$$

it is important include an evaluation of the stiffness of the structure in the optimization process.

To do this, many different authors introduced the idea of stiffness characterization, in many different fashions. Nishiwaki [88] introduced the following objective function

$$\frac{MSE}{\omega \cdot SE_1 + (1-\omega)SE_2} \tag{5.28}$$

This objective function results from the study of the Pareto optima, which represent the trade-off of the minimization mutual mean compliance, and the minimization of the compliance, of both input and output ports. Furthermore, even Zhu, Zhang and Fatikow [96] presented a multi-objective formulation

$$-u_{OUT} + \alpha C_{IN} + \beta C_{OUT} \tag{5.29}$$

As alternative approaches, here are reported two researches that focus on the evaluation of the deformation of the continuum in order to define an effective objective function: the first one is the work by Lee and Gea [97], who adopted the minimization of the sum of the effective strain of the elements, combined with the maximization of the mutual strain energy. On the other hand, particularly remarkable is the work of Yin and Ananthasuresh, [98], disclosing an optimal criterion based on the distortion of the elements: based in this idea, it is possible to obtain a topology for a compliant mechanism in which the deformation is actually distributed over all the structure. A resume of all the cited papers and relative objective functions, functional models, adopted approaches, optimization algorithms and outcomes is reported in appendix D.

Despite the great interest in field, anyway, at the present days, there is not an accepted standard procedure, and the topic of continuum topology optimization for the design of compliant mechanism is still an open issue. For this reason, the adoption of the so-called discrete approaches is an appealing alternative. Anyway, before introducing the discrete approaches to the design of the compliant mechanisms, two researches are reported below that represent an ideal bridge between the continuum approach and discrete approaches.

Characteristic stiffness

Wang and Chen [99] proposed the adoption of the characteristic stiffness approach. Generalizing what has been said in the previous section, in the case of multiple ports, subject to multiple loads, it is possible to arrange the stiffness matrix, the vector of the nodal displacements, and vector of external loads as follows:

$$\begin{vmatrix} k_{1,1} & \cdots & k_{1,m} & k_{1,m+1} & \cdots & k_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ k_{1,m} & \cdots & k_{m,m} & k_{m,m+1} & \cdots & k_{m,n} \\ k_{1,m+1} & \cdots & k_{m,m+1} & k_{m+1,m+1} & \cdots & k_{m+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ k_{1,n} & \cdots & k_{m,n} & k_{m+1,n} & \cdots & k_{n,n} \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ \vdots \\ u_m \\ u_{m+1} \\ \vdots \\ u_n \end{vmatrix} = \begin{vmatrix} f_1 \\ \vdots \\ f_m \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \tag{5.30}$$

$$= \begin{Vmatrix} \underline{\underline{K}}_{11} & \vdots & \underline{\underline{K}}_{12} \\ \cdots & \cdots & \cdots \\ \underline{\underline{K}}_{12} & \vdots & \underline{\underline{K}}_{22} \end{Vmatrix} \cdot \begin{Vmatrix} \underline{U}_1 \\ \cdots \\ \underline{U}_2 \end{Vmatrix} = \begin{Vmatrix} \underline{F}_1 \\ \cdots \\ \underline{0} \end{Vmatrix}$$

Similarly to what has been done in the case of the single input single output case, it is possible to define the matrix which relates only the nodes subject to loads, to the applied forces. In fact, being

$$\begin{cases} \underline{\underline{K}}_{11} \cdot \underline{U}_1 + \underline{\underline{K}}_{12} \cdot \underline{U}_2 = \underline{F}_1 \\ \underline{\underline{K}}_{12} \cdot \underline{U}_1 + \underline{\underline{K}}_{22} \cdot \underline{U}_2 = \underline{0} \end{cases} \tag{5.31}$$

form the second equation it is possible to calculate $\underline{U}_2$, the vector of the nodes that we do not want to include in the model

$$\underline{U}_2 = -\underline{\underline{K}}_{22}{}^{-1} \cdot \underline{\underline{K}}_{12} \cdot \underline{U}_1 \tag{5.32}$$

getting the following by substituting in the first equation of the system:

$$\left(\underline{\underline{K}}_{11} - \underline{\underline{K}}_{12}{}^{T} \cdot \underline{\underline{K}}_{22}{}^{-1} \cdot \underline{\underline{K}}_{12}\right) \cdot \underline{U}_1 = \underline{\underline{K}}_{E} \cdot \underline{U}_1 = \underline{F}_1 \tag{5.33}$$

The most important remark about this formulation is that, differently by the one based on the strain energy and the mutual strain energy, an optimization procedure based on the evaluation of the matrix $\underline{\underline{K}}_{E}$ does not depend by the particular load condition and the configuration of the system (nodal displacements). Based on this observation, Chen and Wang [100], as it has been already showed, proposed the following objective function

$$e^{-(GA-GA^*)} \cdot K_i \cdot K_o \tag{5.34}$$

where GA is the geometric advantage, GA^* the objective geometric advantage, and K_i and K_o the input port and output port characteristic stiffness, considering only the components along the input force and output force directions (for this reason, K_i and K_o are scalar quantities).

Kinetic-static approach

The adoption of the characteristic stiffness for the definition of the objective function allows to consider, in the optimization process, the intrinsic stiffness characteristics of the structure. On the other hand, it may be noticed that, by definition, GA takes in account only the output displacement along the output force direction, and the input displacement

along the input force direction; in this way, there is not any prescription about the displacements along the other directions, which should be neglected for instance.

An effort to overcome this aspect, which is typical of most of the approaches reported in literature, has been done by Wang [83]. In fact, he proposed a kinetic-static (KS) approach, with the purpose of incorporating in the synthesis of the design of the compliant mechanism, not only the prescription of the displacement of the output port in a certain direction, but even in all the other directions, and for a complete set of possible load conditions.

The reasoning starts with the integration, in the characteristic stiffness model, as introduced in the previous section, of a further model for the description of rigid body motions for the input and output forces. To do this, it is possible to define two different sets of nodes rigidly linked each other. Let us start by writing an augmented version of the previously introduced stiffness model

$$\begin{vmatrix} \underline{\underline{H_i}} & \underline{\underline{H_s}} \\ \underline{\underline{H_s}} & \underline{\underline{H_o}} \end{vmatrix} \cdot \begin{vmatrix} \underline{u_i} \\ \underline{u_o} \end{vmatrix} = \begin{vmatrix} \underline{f_i} \\ \underline{f_o} \end{vmatrix} \tag{5.35}$$

where the symbols have the usual meaning, excepted for the dimensions. In fact, in this case, the input and output port are not single degrees of freedom of single nodes, but correspond to a set of nodes $\underline{u_i}$ and $\underline{u_o}$. The nodes of the input and output port are rigidly linked, which means that it is possible to describe their displacement only considering the displacement and the rotation of a unique node for the input port, and a unique node for the output port. This is a fundamental feature of this approach because it makes possible to consider rigid motion modes, and consider both rigid and flexible links, similarly to what will be shown in the next section.

Consider the mechanism depicted in figure 5.6: the set of the nodes of the input port are highlighted in red, and, according to the finite elements framework, in a plane problem, adopting isoparametric elements, their configuration is described by eight Lagrange variables; nevertheless, if these nodes are rigidly connected, so that the element has infinite rigidity, and it is possible to describe their configuration using three parameters, collected in the vector $\underline{\Delta_i}$

$$\underline{u_i} = \begin{vmatrix} \underline{\underline{I}} & \underline{\underline{S}}_{i1} \\ \underline{\underline{I}} & \underline{\underline{S}}_{i2} \\ \vdots & \vdots \\ \underline{\underline{I}} & \underline{\underline{S}}_{ih} \end{vmatrix} \cdot \left| \begin{matrix} \underline{\Delta x_i} \\ \underline{\Delta \theta_i} \end{matrix} \right| = \underline{\underline{T_i}} \cdot \underline{\Delta_i} \tag{5.36}$$

where, $\underline{\underline{S}}_{ik}$ is the rotation skew-symmetric matrix for the node k-th node. Similarly, it can be provided a matrix relation for forces and torques applied to the input port:

$$\underline{\underline{T_i}}^T \cdot \underline{f_i} = \underline{P_i} \tag{5.37}$$

The same matrix equations can be provided for the output port, as a relation between nodal displacements $\underline{u_o}$ (the set of degrees of freedom of all the nodes in the output port set), and the displacement and rotation of the reference node $\underline{\Delta_o}$:

$$\underline{u_o} = \begin{vmatrix} \underline{\underline{I}} & \underline{\underline{S}}_{o1} \\ \underline{\underline{I}} & \underline{\underline{S}}_{o2} \\ \vdots & \vdots \\ \underline{\underline{I}} & \underline{\underline{S}}_{ok} \end{vmatrix} \cdot \left| \begin{matrix} \underline{\Delta x_o} \\ \underline{\Delta \theta_o} \end{matrix} \right| = \underline{\underline{T_o}} \cdot \underline{\Delta_o} \tag{5.38}$$

Furthermore, the correspondent relation for forces and torques can be provided:

$$\underline{\underline{T_o}}^T \cdot \underline{f_o} = \underline{P_o} \tag{5.39}$$

Using the transformation matrixes $\underline{\underline{T_i}}$ and $\underline{\underline{T_o}}$, it is possible to obtain the stiffness model of the compliant mechanism which takes in account the set of rigidly connected nodes at the input and output port

$$\underline{\underline{R}} \cdot \left| \begin{matrix} \underline{\Delta_i} \\ \underline{\Delta_o} \end{matrix} \right| = \begin{vmatrix} \underline{\underline{R_i}} & \underline{\underline{R_s}} \\ \underline{\underline{R_s}} & \underline{\underline{R_o}} \end{vmatrix} \cdot \left| \begin{matrix} \underline{\Delta_i} \\ \underline{\Delta_o} \end{matrix} \right| = \left| \begin{matrix} \underline{P_i} \\ \underline{P_o} \end{matrix} \right| \tag{5.40}$$

where

$$\begin{cases} \underline{R_i} = \underline{\underline{T_i}}^T \cdot \underline{\underline{H_i}} \cdot \underline{\underline{T_i}} \\ \underline{R_s} = \underline{\underline{T_i}}^T \cdot \underline{\underline{H_s}} \cdot \underline{\underline{T_o}} \\ \underline{R_o} = \underline{\underline{T_o}}^T \cdot \underline{\underline{H_o}} \cdot \underline{\underline{T_o}} \end{cases} \tag{5.41}$$

The matrix $\underline{\underline{R}}$ can be referred as mechanism stiffness matrix. This formulation is particularly interesting because it allows to implement in the continuum optimization process, not only taking in account the prescribed desired displacements, but considering specific motion modes which must be avoided as well. Analytically, this is done imposing

some conditions on the eigenvalues of the stiffness matrix. In this way, it is possible to avoid parasitic motions for the compliant mechanism [83].

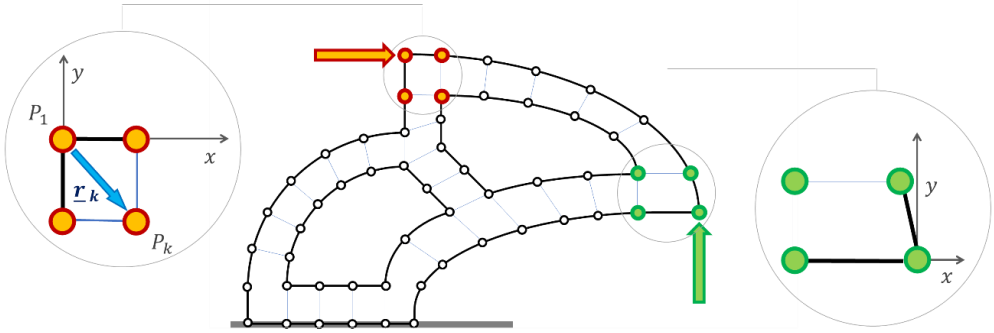


Figure 5.6: sets of nodes at input port and output port

This procedure aims to lead to an optimal structure, provided of a stiffness matrix which fits the functional requirements, comprising all the possible load cases, by the mean of a topology optimization process. In the next section it will be introduced an alternative methodology for the synthesis of the compliant mechanism based on the definition of the stiffness model of the mechanism by the definition of the matrix $\underline{\underline{R}}$, but it will be done by the mean of a strategy based on the multilevel hierarchical paradigm introduced in chapter 2.

5.2.7 Convexity of the optimization problem

A last remark about the synthesis of continuum compliant mechanism using topology optimization procedures regards the convexity of the optimization problem. In chapter 3, it had been stated that one of the issues to address dealing with optimization problem is the convexity of objective functions and constraints, conditions which may ensure to reach a global solution while adopting gradient-based optimization procedures.

As it has been already highlighted, in structural topology optimization, the adoption of the minimization of the mean compliance criteria leads to a convex optimization problem; on the other hand, Lau et al. [101] studied the convexity of the optimization problem adopting the geometric and mechanical advantage as objective functions, and showed

different examples of continuum and ground structures, in which actually the problem is not convex.

5.3 Discrete approach to the compliant mechanism design

Continuum topology optimization is a promising tool for the effective design of compliant mechanisms, but, as it had been shown in the previous section, suffers of some drawbacks such as the non-convexity of the optimization problem, and the dependence by the particular task model adopted. In order to overcome these issues, the adoption of the characteristic stiffness approach has been proposed; nevertheless, as a possible alternative, it may be suggested the use of the so-called discrete approach.

As it had been stated at the beginning of this chapter, Howell disclosed the Pseudo Rigid Body Model (PRBM) [84]. Among other purposes, this approach has been used in order to study different classes of compliant mechanisms, and in particular the Lamina Emergent Mechanisms (LEMs) [102][103][104]. In these works, LEMs has been studied adopting both closed form solution for the linear elastic deformation, and concentrated parameters representation.

Furthermore, beside the investigation of topology optimization of continuum structures, Sigmund et al. approached the problem of the optimization of the compliant structure under point of view of mechanism synthesis [105]; among other considerations, this work is interesting because it highlights the possibility to represent the topology of the structure with the use of graph theory. Furthermore, Quennouelle and Gosselin disclosed a general kineto-static model for compliant parallel manipulators, based on kinematic constraints, and the static equilibrium equations [106]. Energy methods for structural analysis have been investigated by Chen et al. [107], in order to provide an alternative analysis methodology.

The research of Ling et al. [108] is particularly relevant because it deals with many key aspects of the synthesis of compliant mechanism: the proposed approach combines KS model for the characterization of serial and parallel sub-mechanisms, with a representation method based on a condensation of the subsystems composing the main structure. As a result, the method allows to simplify the compliant structure in a simple two node equivalent system.

Since the discrete models appear to be an interesting alternative to continuum topology optimization, in the next subsection, a novel discrete approach for the synthesis of compliant mechanisms will be shown, whose outcomes have been presented in an international conference paper [109].

5.4 Optimal synthesis of topology for compliant mechanisms

At the begin of the present chapter, it had been stated that compliant mechanism are particular devices having characteristics of both mechanisms and structures.

As recalled in chapter 4, mechanisms are generally (and ideally) composed by a set of rigid links connected by the means of actuated kinematic couples. The configuration of a mechanism is identified by the value of a set of variables, able to completely describe the state of the system, and, moreover, according to the Lagrange Equations, the laws of motion of the system may be described by the description of the evolution of such variables in time. In general, we refer to these variables as generalized or Lagrange variables.

As an example, it had been shown that a robotic manipulator is a device having, as a primary goal, the placement of a terminal link, according to a prescribed position and pose. In order to model this kind of systems, for both analysis and control purpose, the primary task is translated to the problem of defining the transformation matrix from a ground reference system, to a reference system attached to the terminal link. For this purpose, Denavit and Hartenberg provided the conventional expressions which describes the transformation of a reference system to the other, in function of the geometry of the links, and the nature of the joints connecting the links. Kinematic, static and dynamic relations may be easily derived using such convention. As an important remark, it may be highlighted that, for the description of rigid links mechanism, even if equilibrium of the forces and congruence, in the form of kinematic constraints, must be ensured, the elastic constitutive equations are not provided in general.

This is not true anymore if the compliant mechanism is modelled as a structure: more in detail, here a structure is meant to be a continuous system, which occupies different regions of Euclidean space, according to the forces that are applied to it. When a force system is applied to a deformable body, the unloaded configuration and the loaded

configuration may be put in relation referring to a displacement field, as it had been shown in chapter 3. As a consequence, unlike standard mechanisms, compliant mechanisms are able to store elastic energy in their structures.

There are different approaches for the analysis of continuum deformable structures: closed form solutions are always the most desirable option, but usually they are difficult to be obtained; the most common option is the finite element method, which is widely used. Nevertheless, for the study of compliant mechanisms, Howell proposed the use of the pseudo rigid body model (PRBM), a model with concentrated parameters, in which an equivalent rigid body system is used to approximate the elastic element. The method is based on the minimization of the error in calculating the position of the extremity of the beam as the pose of the reference system bound on the extreme point of the rigid member. Expressing the position of this reference system will be one of the key points of the methodology proposed in this dissertation.

Reassuming, in the analysis, and eventually in the synthesis, of the compliant mechanisms, there are two points of view to be taken in account: firstly, there is the mechanism point of view, which focuses on the description of the kinematic aspects, describing parameters related to the relative motion of the parts; on the other hand, there is the structure perspective, which mainly studies the response of the system to the external forces, highlighting the relation between mechanical stress and deformation. For this reason, it will be presented here a method for the analysis and synthesis of compliant mechanisms based on an ontological framework in which both mechanism and structures models are represented, and in which the idea of hierarchical organization of the structures is implemented, as it is suggested by the biomimetic studies.

5.4.1 Ontology requirements and taxonomy

Under a philosophical point of view, an ontology may be defined as an “explicit specification of a conceptualization”, or as a “formal specification of a shared conceptualization” [110]. In order to deal with technical tasks such as the analysis of compliant mechanisms, it is necessary to adopt a more specific definition, used in the field of informatics and computational disciplines: an ontology is the definition of a

taxonomy of elements, and the specification of the relations between these elements [111].

A feasible schema for the creation of reliable models for compliant mechanisms must have some specific features:

- it should be able to describe the kinematic of a generic point of the mechanism;
- it should allow to describe in a univocal way the functional requirements: as it will be shown later on, this will be done with the definition of a set of degrees of freedom granted to the reference system of the end effector, or a prescribed rigidity of the system;
- it is possible to decompose the whole system in sub-systems which may be considered as singular elements, or, in other words, to define a hierarchical arrangement of the elements: every sub-structure may be replaced by an equivalent component, which behaves according to the characteristic of the elements of the substructure, and its topology. In other words, it is possible to define different subset of elements, which have a well-known behaviour, and may be considered as black boxes.

At the base of an otology there is the taxonomy of its elements, which represent the hierarchical organization of its classes. For compliant mechanisms, the classes represent the main conceptual entities required for their complete description, both under the mechanism and structure point of view. The complete tree of the taxonomy is depicted in figure5.7, and all its elements will be briefly described in the following.

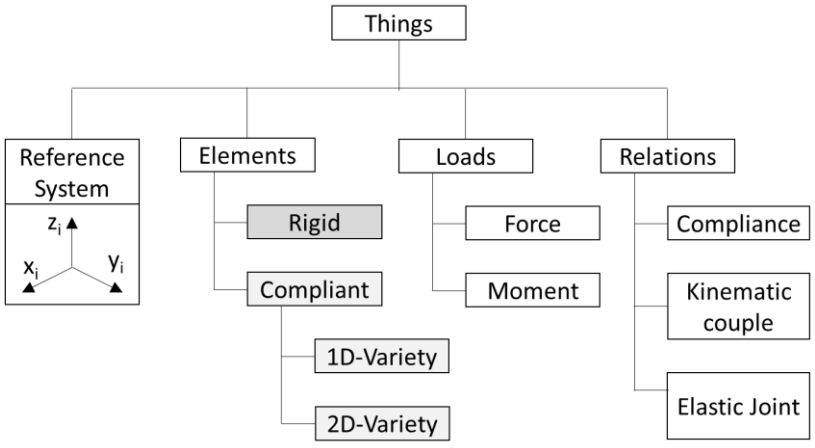


Figure 5.7: elements taxonomy of the ontology for the synthesis of compliant mechanisms

Reference systems and applied forces

At the beginning of this chapter, it had been introduced the concept of functional requirement for the compliant mechanism in terms of displacements of the input and output ports, in the case of geometric advantage, or loads applied to input and output ports, in the case of the mechanical advantage. Actually, if we want to model a flexible system, it is necessary describe the position of points and loads. For this reason, the first two classes of the taxonomy are the reference frames and the applied forces. As we will see later on, if a force is applied on an element of the mechanism, the application point will be identified by the origin of a reference frame. This is important in order to describe the contribution of the force to the total potential energy as the work done by the force for a feasible virtual displacement.

As an important remark, it is convenient here highlight the fact that both forces and torques can be considered as loads, so it is possible to say that are generalized loads, similarly to the generalized Lagrange coordinates, both linear and angular.

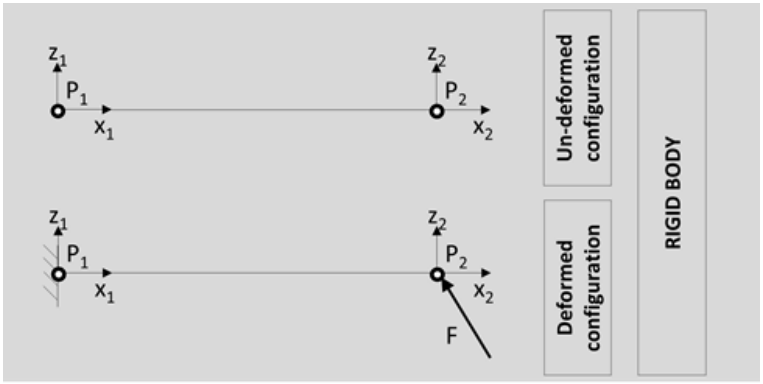
Rigid and compliant elements

A further taxonomy class represents the elements which actually compose the structure of the compliant mechanism. These basic components can be of two kinds: perfectly rigid links, or flexible elements:

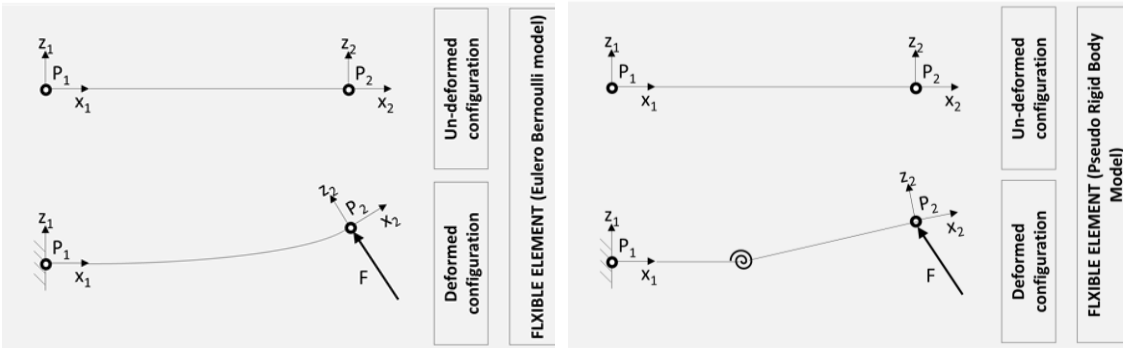
- a rigid link is characterized by the fact that the position and the orientation of all its possible attached reference systems depends only by six parameters (considering the 3D space, three parameters in the plane case). In other words, if it is known the position and the pose of a reference frame on the rigid element, consequently it is known the pose and position of any other reference frame attached to it, and, furthermore, this spatial relation depends only by the relative position of the frames, and does not depends by the forces applied on the rigid body;
- differently by rigid elements, flexible elements occupy different regions of space depending on the loads to which they are subject. As a convention, the region of space occupied by the element while there are no loads applied on, is called un-deformed configuration; all the other cases, when there are forces and moments applied, are deformed configurations. As a further subdivision, it is possible distinguish between two further types of flexible elements:
    - 1D variety elements: actually, these are beams and trusses elements;

- 2D variety elements: plates and shells;

The main difference between rigid and flexible elements is actually the way to determine the position and pose of a generic attached reference system with to respect a principal reference frame: for a rigid body this is simple, but, for flexible elements, it must be done in accordance to the deformation state of the continuum body. In other words, in order to determine the pose of a generic reference system, it is necessary to solve the elastic structure. This may be done using different methods: as an example, it may be used the finite elements method, or, alternately, it may be adopted a concentrated parameter model, using the Lagrange Equations.



(a)



(b) *(c)*

Figure 5.8: models of a rigid rod (a), a flexible beam according to the Eulero Bernulli model (b), a flexible beam according to the concentrated parameters model (c)

Figure 5.8 depicts the differences of the models for three different kinds of elements: figure 5.8(a) shows the model for a rigid rod, figure 5.8 (b) shows the behaviour of a

flexible beam, adopting the Eulero Bernoulli model, and figure 5.8(c) depicts the representation of a flexible beam considering the pseudo rigid body model. It is important to highlight that any model is a good candidate in order to be implemented in the framework, as long as it is able to describe the mechanical response of the element to applied loads, and determine the elastic energy in function of the displacements.

Links between elements

Once structural elements and the forces have been defined, the next step is describing the way they interact each other. In the present research, three fundamental modes of interaction are reported: congruence, kinematic coupling, and elastic coupling. More precisely, all these three relations involve always a couple of reference systems, and every reference system belongs to a different element or its origin is the application point generalized load.

The simplest relation is the congruence, that basically impose the two reference systems to be coincident. In general, this relation is used in order to connect a rigid element and a flexible element.

The second relation is the kinematic couple, which imposes the existence of one or more relative degrees of freedom between the two reference frames. As introduced in the previous chapter, the most common way to represent this relation is the use of the transformation matrices, which allows to change the representation of a vector, passing from a reference system to another. As an example, the matrix

$$\boldsymbol{R}\,(\theta, z) = \begin{bmatrix} cos\theta & -sin\theta & 0 & 0 \\ sin\theta & cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.42}$$

models the relative rotation around the z axis of an angle θ.

The third relation is the elastic coupling, and, as the case of the kinematic couple, it allows the relative movement of the two reference frames. But differently from the previous case, the relative (angular or linear) displacement is dependent by the loads applied on the elements the reference systems are attached to. The relation between the loads and the displacement can be, for instance, linear

$$\begin{aligned} \underline{F} &= k_{lin} \cdot \underline{s} \quad && Traslation \\ \underline{\tau} &= k_{rot} \cdot \underline{\theta} \quad && Rotation \end{aligned} \tag{5.43}$$

indicating a linear elastic behaviour. In this particular case, the work of such elastic forces represents the elastic energy stored, and will be taken in account writing the Lagrange equations.

Congruence, cinematic coupling and elastic coupling may be represented as transformation matrices, which are functions of the Lagrange variables. This means that, following the kinematic chain of the CM, it is always possible to relate vector quantities, such as forces and displacements, in function of the generalized coordinates.

5.4.2 Analysis of the compliant mechanism

Representation of the compliant mechanism model

In order to provide an example of the application of the proposed taxonomy for the representation and analysis of the CM, it will be now considered a well-known compliant structure, belonging to the class of LEM mechanism: the simple area reduced joint.

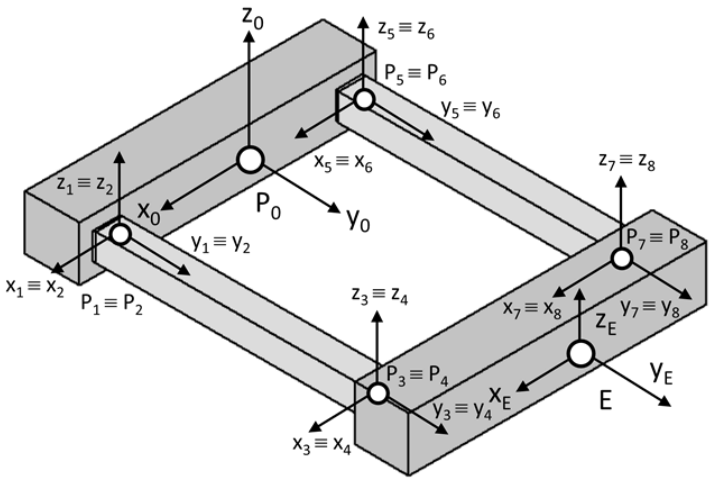


Figure 5.9: representation of the simple area reduced joint

This particular component is shown in figure 5.9, and it can be briefly described as a structure composed by two rigid elements, one grounded and one free, connected each other by two flexible beams. Furthermore, there are showed the reference frames attached to the rigid and flexible elements; in particular, the reference system $P_0 x_0 y_0 z_0$ is considered grounded, meanwhile the $P_E x_E y_E z_E$ the reference system in correspondence

of the application point of the loads. Furthermore, figure 5.10 shows a block diagram representing the simple area reduced joint, in which every block represents an element or a relation, according to the proposed taxonomy.

It may be found in literature that this kind of structure can be adopted in order to realize a compliant hinge: this means that the functional requirement for this particular structure is allowing the rotation of the rigid element E around the axis x_E, and neglect rotation and displacements in the other directions. But how can be formalize this particular task?

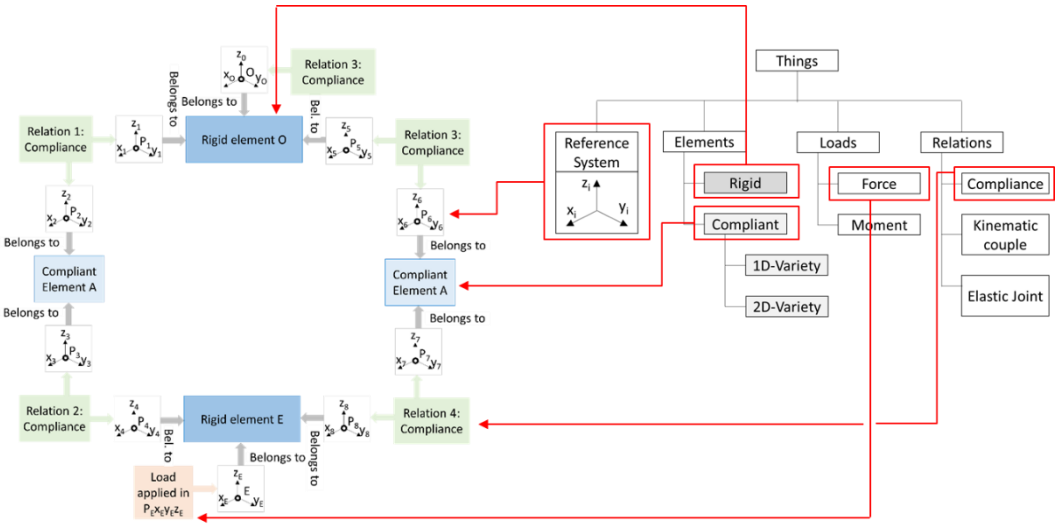


Figure 5.10: block diagram representing the simple area reduced joint and corresponding elements of the taxonomy

Definition of the subset of the rigidity matrix elements

The task of realizing a compliant joint may better defined as the requirement for the structure to allow certain degrees of freedom, and deny others. In other words, the system should have a low stiffness in correspondence of a sub-set of generalized or Cartesian coordinates, and high stiffness for the others.

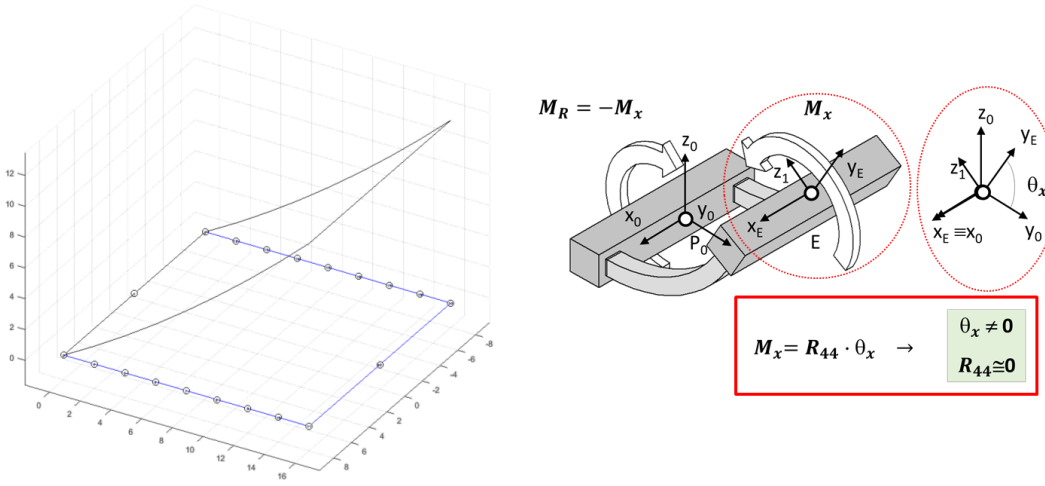


Figure 5.11: example of tortional mode of deformation for the simple area reduced joint

To do this, it is necessary to identify a set of possible generalized loads that can be applied to the structure, and then define, as functional requirement, its mechanical response. This can be done, for instance, by carrying out a finite element analysis of the structure for every load case. As an example, figure 5.11 shows the analysis of the simple area reduced having the element E, subject to a torque M_x, applied along the x_E axis: this particular load produces a deformation of the structure, and in particular a remarkable displacement of the point E along the axis z_E, and a rotation of the reference system $P_E x_E y_E z_E$ along the x_E axis.

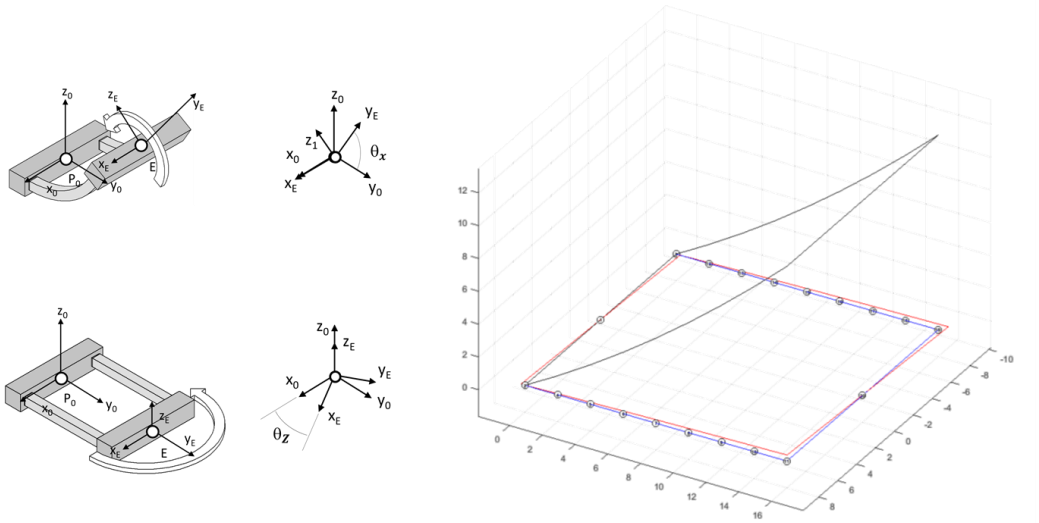


Figure 5.12: comparison between two tortional deformation modes for the simple area reduced joint

This is only one of the possible mechanical responses of the CM, but actually, the main goal is determining which deformation modes are compatible with the functional requirements, which, in this case, are allowing or neglecting certain rotations or displacements.

Figure 5.12 shows the comparison between two deformation modes of the simple area reduced joint: one is corresponding to the application, a dummy torque M_x, applied along the x_E axis; the other one corresponds to the application of a dummy torque M_z, applied along the z_E. As it was expected, the result is that, numerically, the entity of the rotation θ_x around the x_E axis is much higher than the rotation θ_z around the z_E axis.

The better way to condense the information about different deformation modes of a structure is, obviously, the stiffness matrix:

$$\begin{bmatrix} F_x \\ F_y \\ F_z \\ M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & R_{14} & R_{15} & R_{16} \\ R_{21} & R_{22} & R_{23} & R_{24} & R_{25} & R_{26} \\ R_{31} & R_{32} & R_{33} & R_{34} & R_{35} & R_{36} \\ R_{41} & R_{42} & R_{43} & R_{44} & R_{45} & R_{46} \\ R_{51} & R_{52} & R_{53} & R_{54} & R_{55} & R_{56} \\ R_{61} & R_{62} & R_{63} & R_{64} & R_{65} & R_{66} \end{bmatrix} \cdot \begin{bmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \\ \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} \tag{5.44}$$

The results of the previous two simulations, caried out in MATLAB, are in accordance to the fact that the $R_{66} \gg R_{44}$, which correspond to the fact that the structure is much stiffer in a certain direction with to respect the other.

Consequently, it is possible to give a formal definition of the synthesis process: actually, it consists in determining the topology, shape and elements sizes of the structure which behaves providing a prescribed set of elements of the stiff ness matrix.

Ideally, to completely address the functions of the mechanism, it should be necessary to specify all the components of the stiffness matrix. For simplicity here only four components will be taken in account, and figure 5.13 depicts the 4 correspondent load cases.

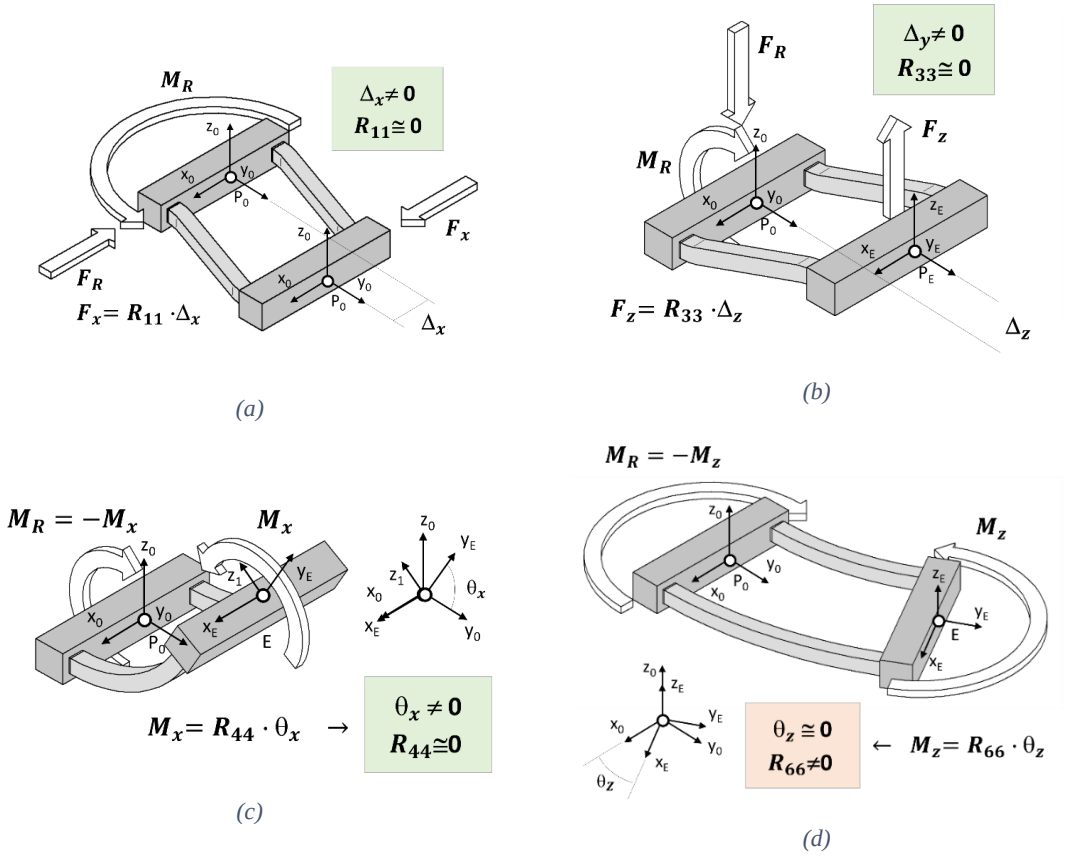


Figure 5.13: deformation modes taken in account in the analysis of the simple area reduced joint

As a consequence of this simplification, it is possible to highlight the elements of the stiffness matrix which are representative of the mechanical responses:

$$\begin{bmatrix} F_x \\ \dots \\ F_z \\ M_x \\ \dots \\ M_z \end{bmatrix} = \begin{bmatrix} R_{11} \cong 0 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & R_{33} \cong 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & R_{44} \cong 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & R_{66} \neq 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta_x \\ \dots \\ \Delta_z \\ \theta_x \\ \dots \\ \theta_z \end{bmatrix} \tag{5.45}$$

Actually, figure 5.13 shows the different deformation modes corresponding to different load conditions. The set of all these responses, according to Wang [83] and Ling et al. [108], is the functional requirement for the mechanism. Furthermore, figure 5.14 shows that the stiffness matrix resulting by the analysis caried out in MATLAB is coherent with the one reported in equation 5.45.

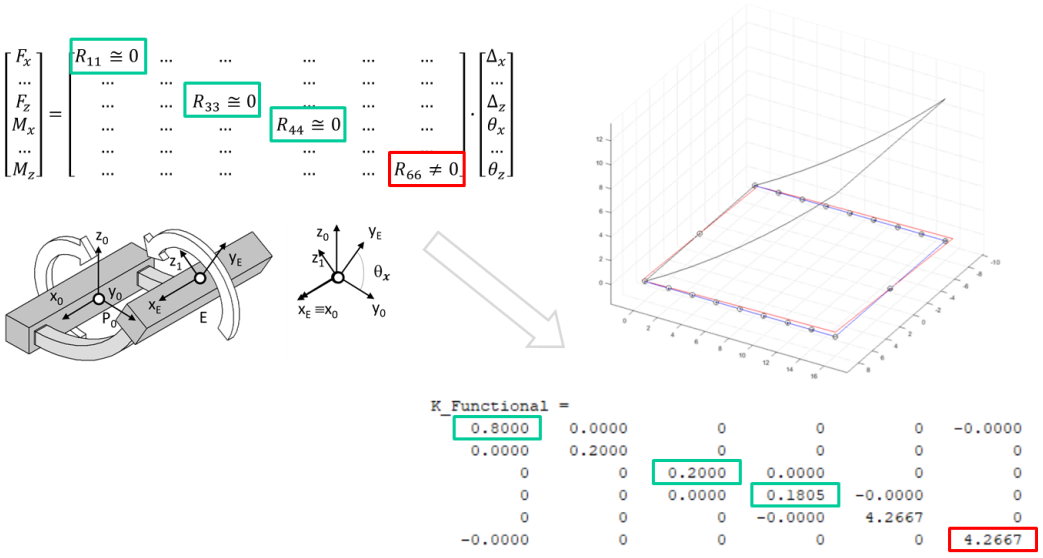


Figure 5.14: stiffness matrix computed during the simulation of the simple area reduced joint

Until now it had been presented a procedure for the analysis of the simple area reduced joint, but what if I want to change this functional requirement? This question leads to the problem of the synthesis of the CM, or, in other words, the identification of a new structure, a new topology of the mechanism based on a different functional requirement. In order to carry out this goal, in the next subsections will be presented a method based on the modify of the morphology of the structure, modelled according to the discrete approach, in the framework of the proposed ontology: in fact, the generation of the new topologies are based on both the adoption the multilevel hierarchical organization of substructures, and the use of different elements being part of the taxonomy.

5.4.3 Mechanism synthesis using hierarchical organization

Let us consider the goal of modifying the response of the simple area reduced joint, so that it may not allow anymore the rotations around the axis x_E, keeping free the shear displacements along the x_E axis and the z_E axis. According to the proposed methodology, such behaviour may have expressed by the matrix equation:

$$\begin{bmatrix} F_x \\ \dots \\ F_z \\ M_x \\ \dots \\ M_z \end{bmatrix} = \begin{bmatrix} R_{11} \cong 0 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & R_{33} \cong 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & R_{44} \neq 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & R_{66} \neq 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta_x \\ \dots \\ \Delta_z \\ \theta_x \\ \dots \\ \theta_z \end{bmatrix} \tag{5.46}$$

It is not intuitive to find a structure which should have such mechanical response. Here the proposed approach is the substitution of the constitutive elements of the structure with another type of elements having a different mechanical response. According to the ontological framework introduced earlier, this substitution may be done using or a set of elements belonging the classes of the taxonomy, or, alternately, considering more complex structures having the same input/output configuration. As an example, figure 5.15 shows how it is possible to modify the morphology of the simple area reduced joint by the means of the replacement of its 1D flexible elements with two substructures. In this case, the two substructures are two simple area reduced joint as well: actually, this is an application of the hierarchical organization paradigm.

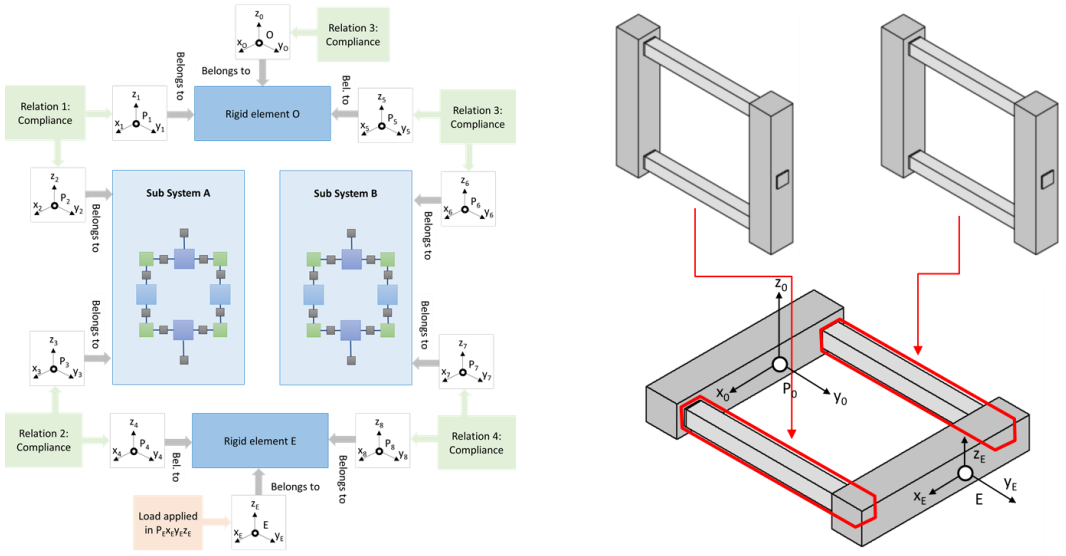


Figure 5.15: example of application of the hierarchical organization of the structures for the synthesis of a compliant mechanism

Figure 5.16 depicts the mechanical response of the structure: as it can be notice, while in the case of the simple area reduced joint the stiffens of the structure to a torque applied to the x_E axis is very small, in this new structure the parameter of the stiffness matrix R_{44}

is higher; on the other hand, the topology change did not affect the other two parameters, R_{11} and R_{66}. Actually, it can be said that the new system fulfils the functional requirements.

As an important remark, it must be highlighted that the new behaviour of the structure had been obtained using the same structural elements of the original arrangements, which means that the mechanical response does not depend by the characterization of the elements, but only by the way these elements are arranged. This consideration is coherent with the observation of living beings: in fact, it is well known that animals and plants are able to synthetize tissues having a wide range of different performances, using few basic structural elements arranged according to a great number of different configurations.

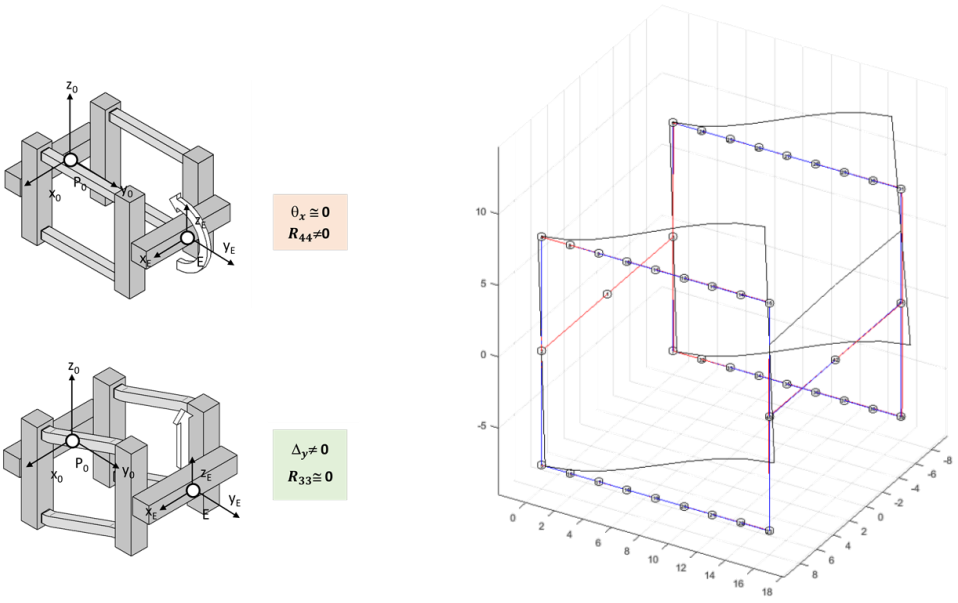


Figure 5.16: two different deformation modes of the hierarchically modified simple area reduced joint

5.4.4 Mechanism synthesis modelling the constitutive elements

Let's assume now that we want to modify further the mechanical response of the system according to the following stiffness matrix:

$$\begin{bmatrix} F_x \\ \dots \\ F_z \\ M_x \\ \dots \\ M_z \end{bmatrix} = \begin{bmatrix} R_{11} \cong 0 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & R_{33} \neq 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & R_{44} \neq 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & R_{66} \neq 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta_x \\ \dots \\ \Delta_z \\ \theta_x \\ \dots \\ \theta_z \end{bmatrix} \tag{5.47}$$

This particular set of rigidity parameters correspond to a high rigidity for all the load conditions except the force applied along the direction x_E.

This time, instead of resorting to the concept of hierarchical organization, on the contrary it can be adopted the use of elements having a different mechanical characterization. Figure 5.17 shows that the compliant elements of the simple area reduced joint, which are represented as a 1D segment, and characterized by the same moment of inertia in all directions, are replaced by other elements, represented as bidimensional varieties, in this case two plates. Plates are characterized by a high moment of inertia in one direction, compared to the one in the normal direction, resulting in a higher angular deflection along one axis, with to respect the perpendicular one.

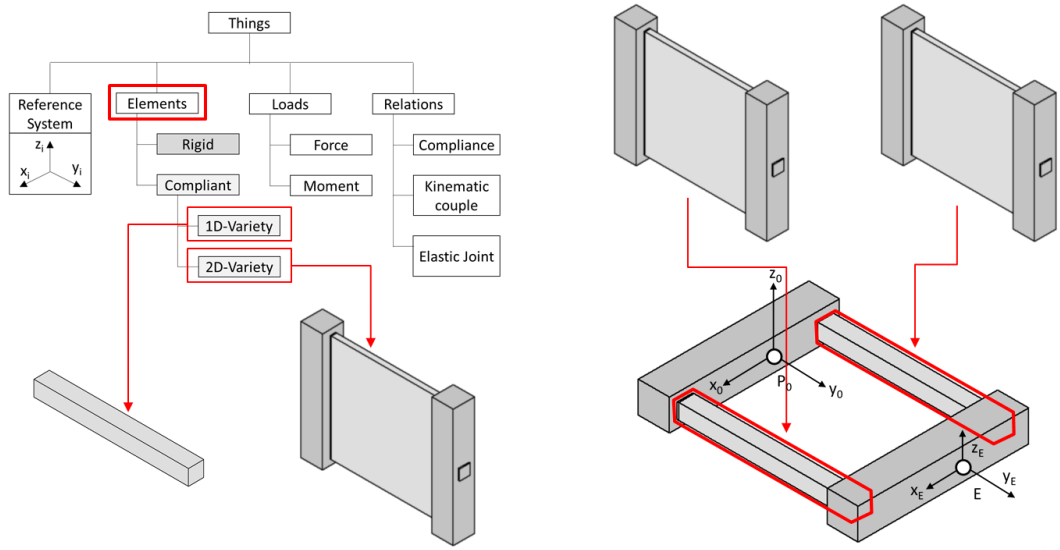


Figure 5.17: example of element type substitution for the synthesis of a compliant mechanism

A representation of this outcome is showed in figure 5.18: the mechanical deformation due to a torque applied along the x_E axis is actually negligible, and the only relevant displacement (both linear and angular) of the point P_E is the linear displacement along the x_E axis.

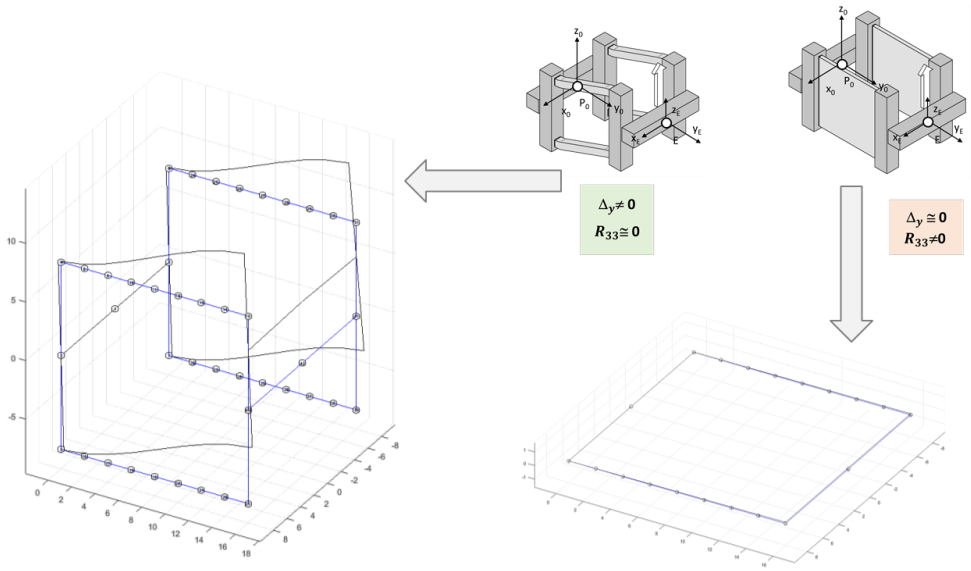


Figure 5.18: confront between the mechanical responses of the two modified simple area reduced joint mechanisms

5.4.5 Use of hierarchical composition of structures as a tool for the design

The examples reported in the previous subsections suggest how it is possible to implement the multilevel paradigm in order to improve the flexibility of the discrete approaches for the synthesis of CM. Actually, the main drawback of methodologies, like the PRBM, is the difficulty in determining new topologies, since the connectivity of the elements is in general prescribed a priori by the designer.

This limitation is typical of size and shape optimizations, and can be overcome only finding a way to introduce or remove new elements in the system. Removing elements is straight forward since it is possible to neglect the influence of an element on the system by imposing an infinitesimal value for a certain distributed quantity, such as the density or the Young modulus. The automated introduction of new elements is more difficult, and this is the reason why optimization of ground structures and continuum topology optimization are based on the determination of a finite and constant set of elements (set of beams or mesh elements of the design space).

What has been discussed so far can be an inspiration for a strategy which aims to enrich the topology of a certain structure, instead of erasing some parts. More in detail, the ontology of multilevel described in chapter 2 postulates the equivalence of elements and

structures: this means that it is possible replace a certain element with a sub-structure which is equivalent in terms of functional response. This implication is particularly useful when a simple element with the desired response is not available, because it suggests the idea of find a sub-structure that fulfils that particular requirement.

Actually, this is the process involved in the synthesis of the modified area reduced joint, described in figure 5.15, and which fulfils the functional requirement expressed by the stiffness matrix in the equation 5.46. In fact, the flexible beam elements are replaced with the area reduced joint sub-strictures because they are not able to offer different moments of inertia in correspondence of different (not axial) directions.

Generalizing, it can be said that design tool based on the hierarchical organization of the structures, should be able to replace the elements of some basic structures with some sub-structures characterized by a specific behaviour, in order to obtain the required functional response of the whole system. This implies that, starting from a little library of basic elements and topologies, it is possible to generate procedurally a great number of final structures.

5.5 Conclusions and future developments

This chapter has been devoted to the study of the implementation of the biomimetic concept of flexibility in the design activity. At the state of the art, this concept has been already largely applied in the field of CM synthesis. The analysis of the literature highlighted that two main approaches are adopted by researchers: the design of CM by the means of continuum topology optimization, and discrete models.

Continuum TO for the design of CM has been largely investigated, and generated a great number of different methods, based on the formulation of different objective functions, and assuming different task models. Nevertheless, continuum CM design shows to be affected by some intrinsic issues, such as difficulty in generating structures with distributed compliance, and not convexity of the optimization problem. On the other hand, the adoption of a discrete approach seems to be convenient in order to overcome such issues, even if, in general, it is difficult to automatically generate new concepts and topologies.

Therefore, in order to overcome the issues of the already existing approaches, it had been proposed the use of a novel ontological framework suitable for the synthesis of CM. The proposed ontology implements the biomimetic concepts of hierarchical organization of the structures introduced in chapter 1, allowing the possibility of generate topologies in a modular way. This kind of procedure represents an improvement with respect to the state of the art, because, on one hand, it is not affected by the difficulties of continuum CM synthesis procedures in realizing distributed compliance, and, on the other hand, overcomes the difficulties in generating new topologies of the discrete approach.

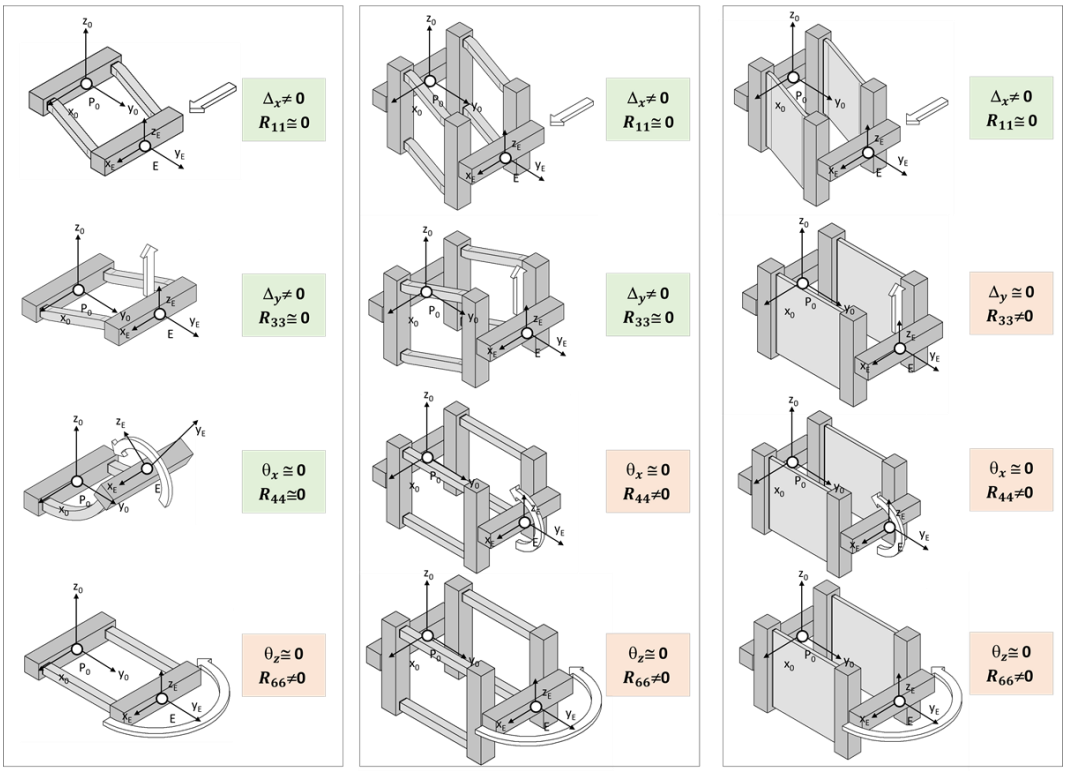


Figure 5.19: examples of compliant mechanisms characterized by different mechanical responses, and, consequently, by different stiffness matrixes

Figure 5.19 depicts the outcomes of the application of the described procedures: starting from a functional requirement, expressed in terms of relative values of some elements of the stiffness matrix of the structure, it is possible to obtain the topologies fulfilling such specifications.

Despite the potentialities of the proposed approach, differently by the outcomes of the chapter 3 and 4, the presented results are mostly conceptual: the ontological framework

is a starting point for the design of a synthesis tool, and the idea of hierarchical organization of the structures and anisotropy of the behaviour of elements have to be formalized and integrated in actual algorithms, and, finally, in numerical procedures. For this reason, the knowledge developed in the study of continuum topology optimization is valuable, and should be inspiration of a further stage of the present research.

5.6 Bibliography

[83] Wang, M. Y. (2008, January). A kinetoelastic approach to continuum compliant mechanism optimization. In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 43260, pp. 183-195).

[84] Howell, L. L. (2013). Compliant mechanisms. In 21st century kinematics (pp. 189-216). Springer, London …

[85] Sigmund, O. (1997). On the design of compliant mechanisms using topology optimization. Journal of Structural Mechanics, 25(4), 493-524.

[86] Lau, G. K., Du, H., & Lim, M. K. (2001). Use of functional specifications as objective functions in topological optimization of compliant mechanism. Computer methods in applied mechanics and engineering, 190(34), 4421-4433.

[87] Salamon, B. A., & Midha, A. (1998). An introduction to mechanical advantage in compliant mechanisms.

[88] Nishiwaki, S., Min, S., Yoo, J., & Kikuchi, N. (2001). Optimal structural design considering flexibility. Computer methods in applied mechanics and engineering, 190(34), 4457-4504.

[89] Wang, M. Y. (2009). Mechanical and geometric advantages in compliant mechanism optimization. Frontiers of Mechanical Engineering in China, 4(3), 229-241.

[90] Dirksen, F., & Lammering, R. (2011). On mechanical properties of planar flexure hinges of compliant mechanisms. Mechanical Sciences, 2(1), 109-117.

[91] Ansola, R., Veguería, E., Maturana, A., & Canales, J. (2010). 3D compliant mechanisms synthesis by a finite element addition procedure. Finite Elements in Analysis and Design, 46(9), 760-769.

[92] LI, Y. Topology Optimization of Compliant Mechanics Based on the BESO Method. 2014 (Doctoral dissertation, PhD Thesis. School of Civil-Enviromental and Chemical Engeneering-RMIT University).

[93] Liu, C. H., Huang, G. F., & Chen, T. L. (2017). An evolutionary soft-add topology optimization method for synthesis of compliant mechanisms with maximum output displacement. Journal of Mechanisms and Robotics, 9(5).

[94] Li, Y., Huang, X., Xie, Y. M., & Zhou, S. W. (2014). Evolutionary topology optimization of hinge-free compliant mechanisms. International Journal of Mechanical Sciences, 86, 69-75.

[95] Huang, X., Li, Y., Zhou, S. W., & Xie, Y. M. (2014). Topology optimization of compliant mechanisms with desired structural stiffness. Engineering Structures, 79, 13-21.

[96] Zhu, B., Zhang, X., & Fatikow, S. (2014). A multi-objective method of hinge-free compliant mechanism optimization. Structural and Multidisciplinary Optimization, 49(3), 431-440.

[97] Lee, E., & Gea, H. C. (2014). A strain based topology optimization method for compliant mechanism design. Structural and Multidisciplinary Optimization, 49(2), 199-207.

[98] Yin, L., & Ananthasuresh, G. K. (2003). Design of distributed compliant mechanisms. Mechanics based design of structures and machines, 31(2), 151-179.

[99] Wang, M. Y., & Chen, S. (2009). Compliant mechanism optimization: analysis and design with intrinsic characteristic stiffness. Mechanics Based Design of Structures and Machines, 37(2), 183-200.

[100] Chen, S., & Wang, M. Y. (2007, January). Designing distributed compliant mechanisms with characteristic stiffness. In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 48094, pp. 33-45).

[101] Lau, G. K., Du, H., & Lim, M. K. (2001). Convex analysis for topology optimization of compliant mechanisms. Structural and Multidisciplinary Optimization, 22(4), 284-294.

[102] Jacobsen, J. O., Chen, G., Howell, L. L., & Magleby, S. P. (2009). Lamina emergent torsional (LET) joint. Mechanism and Machine Theory, 44(11), 2098-2109.

[103] Delimont, I. L., Magleby, S. P., & Howell, L. L. (2015). Evaluating compliant hinge geometries for origami-inspired mechanisms. Journal of Mechanisms and Robotics, 7(1).

[104] Jacobsen, J. O., Winder, B. G., Howell, L. L., & Magleby, S. P. (2010). Lamina emergent mechanisms and their basic elements. Journal of Mechanisms and Robotics, 2(1).

[105] Kawamoto, A., Bendsøe, M. P., & Sigmund, O. (2004). Planar articulated mechanism design by graph theoretical enumeration. Structural and Multidisciplinary Optimization, 27(4), 295-299.

[106] Quennouelle, C., & Gosselin, C. (2011). Kinematostatic modeling of compliant parallel mechanisms. Meccanica, 46(1), 155-169.

[107] Chen, G., Ma, F., Bai, R., Magleby, S. P., & Howell, L. L. (2017, August). A framework for energy-based kinetostatic modeling of compliant mechanisms. In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 58172, p. V05AT08A021). American Society of Mechanical Engineers.

[108] Ling, M., Cao, J., Howell, L. L., & Zeng, M. (2018). Kinetostatic modeling of complex compliant mechanisms with serial-parallel substructures: a semi-analytical matrix displacement method. Mechanism and Machine Theory, 125, 169-184.

[109] Caputi, A., & Russo, D. (2019, November). Optimal Synthesis of Topology for Compliant Mechanisms. In ASME International Mechanical Engineering Congress and Exposition (Vol. 83518, p. V014T14A019). American Society of Mechanical Engineers.

[110] Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology?. In Handbook on ontologies (pp. 1-17). Springer, Berlin, Heidelberg.

[111] Arp, R., Smith, B., & Spear, A. D. (2015). Building ontologies with basic formal ontology. Mit Press..

Appendix A: back propagation and adjoint method

The adjoint method is used to find the gradient of a certain objective function $J(\underline{u})$ with respect to a set of control parameters $\underline{x}$, when a relation exists between the state variables

$\underline{u}$ and $\underline{x}$, such that $\underline{u} = \underline{G}(\underline{x})$. It is important to calculate the derivatives of J with respect to $\underline{x}$, without operating the computation of the state variables $\underline{u}$: this problem is relevant especially when $\underline{u}$ and $\underline{x}$ are vectors of high dimension.

The difficulty in applying the adjoint method depends by the type of the functions J and $\underline{G}$, and in the following three different cases will be illustrated.

Linear function

Let us consider the following optimization problem:

$$\begin{aligned} &min \qquad J(\underline{u}) = \underline{c}^T \cdot \underline{u} \\ &s.t. \qquad \underline{\underline{A}} \cdot \underline{u} = \underline{b}(\underline{x}) \end{aligned} \tag{A.1}$$

where J is a performance index of my system, defined by a weight vector $\underline{c}^T$ that is known, and the state variables $\underline{u}$; furthermore, $\underline{u}$ have to satisfy the state equation characterized by the matrix $\underline{\underline{A}}$, which is known, and by the vector $\underline{b}$, which is a known function of the vector of the design variables $\underline{x}$. In this formulation, the real unknowns are the design variables $\underline{x}$, that I want to determinate in order to minimize the function J. The most natural way to solve the problem is the one which involve the solution of the state variables $\underline{u}$, but this strategy is not efficient if the dimension of $\underline{u}$ is high.

For this reason, the goal is writing J as a function of $\underline{x}$, and this can be done by manipulating the state equation:

$$\underline{\underline{A}} \cdot \underline{u} = \underline{b}(\underline{x}) \quad \Longrightarrow \quad \underline{u} = \underline{\underline{A}}^{-1} \cdot \underline{b}(\underline{x}) \tag{A.2}$$

Then, substituting in the expression of the objective function:

$$J(\underline{u}) = \underline{c}^T \cdot \underline{u} \quad \Longrightarrow \quad J = J(\underline{x}) = \underline{c}^T \cdot \underline{\underline{A}}^{-1} \cdot \underline{b}(\underline{x}) = \underline{\lambda}^T \cdot \underline{b}(\underline{x}) \tag{A.3}$$

which represent a function only of the design variables $\underline{x}$, because the vector $\underline{\lambda}$ is known.

It is interesting to notice that, in the implementation of an optimization algorithm, we should be interested in the determination of the gradient of the function J, and actually the sensitivity of J with to respect the design variables reads as follows:

$$\frac{\partial J(\underline{x})}{\partial \underline{x}} = \underline{\lambda}^T \cdot \frac{\partial \underline{b}(\underline{x})}{\partial \underline{x}} \tag{A.4}$$

This equation represents the sensitivity of the objective function to the variations of the design variables $\underline{x}$.

Explicit backpropagation

Let us consider the following optimization problem:

$$\begin{aligned} &min \qquad J = J(\underline{u}) \qquad\qquad \underline{u} = \underline{u}_1, \underline{u}_2, \dots, \underline{u}_n \\ &s.t. \qquad \underline{u}_{i+1} = \underline{G}_i(\underline{u}_i, \underline{x}) \end{aligned} \tag{A.5}$$

where $\underline{u}$ represents a set of discrete states of the system. This situation is typical when the state of the system depends by the evolution of the system itself. Forward Euler method involves a similar recursive formulation, and the same structure may be found in the neural networks as well.

Again, the objective is writing the partial derivatives of the objective function J with respect to the design variables. As a first step, it is possible to linearize the equations:

$$\begin{aligned} &\delta J = \frac{\partial J}{\partial \underline{u}_n} \cdot \delta \underline{u}_n \\ &\delta \underline{u}_n = \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \delta \underline{u}_{n-1} + \frac{\partial \underline{G}_{n-1}}{\partial \underline{x}} \cdot \delta \underline{x} \\ &\delta \underline{u}_{n-1} = \frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \delta \underline{u}_{n-2} + \frac{\partial \underline{G}_{n-2}}{\partial \underline{x}} \cdot \delta \underline{x} \\ &\vdots \\ &\delta \underline{u}_1 = \frac{\partial \underline{G}_0}{\partial \underline{u}_0} \cdot \delta \underline{u}_0 + \frac{\partial \underline{G}_0}{\partial \underline{x}} \cdot \delta \underline{x} \end{aligned} \tag{A.6}$$

being $\delta \underline{u}_0$ null, or a prescribed initial function of $\underline{s}$.

Once the linearization has been caried out, it is possible to start from the last state, and compute the previous states one by one, backpropagating indeed:

$$\begin{aligned} &\delta \underline{u}_n = \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \left(\frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \delta \underline{u}_{n-2} + \frac{\partial \underline{G}_{n-2}}{\partial \underline{x}} \cdot \delta \underline{x} \right) + \frac{\partial \underline{G}_{n-1}}{\partial \underline{s}} \cdot \delta \underline{x} \\ &\delta \underline{u}_n = \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \delta \underline{u}_{n-2} + \left(\frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{x}} + \frac{\partial \underline{G}_{n-1}}{\partial \underline{x}} \right) \cdot \delta \underline{x} \\ &\delta \underline{u}_n = \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \left(\frac{\partial \underline{G}_{n-3}}{\partial \underline{u}_{n-3}} \cdot \delta \underline{u}_{n-3} + \frac{\partial \underline{G}_{n-3}}{\partial \underline{s}} \cdot \delta \underline{x} \right) + \left(\frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{x}} + \frac{\partial \underline{G}_{n-1}}{\partial \underline{x}} \right) \cdot \delta \underline{x} \\ &\delta \underline{u}_n = \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \frac{\partial \underline{G}_{n-3}}{\partial \underline{u}_{n-3}} \cdot \delta \underline{u}_{n-3} + \left(\frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{u}_{n-2}} \cdot \frac{\partial \underline{G}_{n-3}}{\partial \underline{x}} + \frac{\partial \underline{G}_{n-1}}{\partial \underline{u}_{n-1}} \cdot \frac{\partial \underline{G}_{n-2}}{\partial \underline{x}} + \frac{\partial \underline{G}_{n-1}}{\partial \underline{x}} \right) \cdot \delta \underline{x} \\ &\vdots \end{aligned} \tag{A.7}$$

The number n of the iterations required by the analysis may be high, and the final expression of the gradient of the objective function J with to respect the vector of the

design variables $\underline{s}$ can be very high. For this reason, it is convenient adopt an adjoint method as follows. Let us notice that generic linearized state equation can be written as:

$$\begin{aligned} &\delta\underline{u}_{i+1} = \frac{\partial \underline{G}_i}{\partial \underline{u}_i} \cdot \delta\underline{u}_i + \frac{\partial \underline{G}_i}{\partial \underline{x}} \cdot \delta\underline{x} \quad \Longrightarrow \\ &\Longrightarrow \quad \delta\underline{u}_{i+1} - \frac{\partial \underline{G}_i}{\partial \underline{u}_i} \cdot \delta\underline{u}_i - \frac{\partial \underline{G}_i}{\partial \underline{s}} \cdot \delta\underline{x} = \underline{0} \end{aligned} \tag{A.8}$$

This expression may be added to the formulation of the derivative of the objective function, by the mean of a vector $\underline{\lambda}_{i+1}$ such that

$$\delta J = \frac{\partial J}{\partial \underline{u}_n} \cdot \delta\underline{u}_n = \frac{\partial J}{\partial \underline{u}_n} \cdot \delta\underline{u}_n + \underline{\lambda}_{i+1}^T \cdot \left(\delta\underline{u}_{i+1} - \frac{\partial \underline{G}_i}{\partial \underline{u}_i} \cdot \delta\underline{u}_i - \frac{\partial \underline{G}_i}{\partial \underline{x}} \cdot \delta\underline{x} \right) \tag{A.9}$$

This procedure is the same used in the Lagrange multipliers method, because, actually, it augments a little perturbation of the objective function by the addition of the inner product of the vector of multipliers and the variation of the implicit expression of the state equations, which is null, not considering higher level order terms. This is true for the state equations at each step, so the sum of all contributions must be null as well:

$$\delta J = \frac{\partial J}{\partial \underline{u}_n} \cdot \delta\underline{u}_n = \frac{\partial J}{\partial \underline{u}_n} \cdot \delta\underline{u}_n + \sum_{i=0}^{n-1} \underline{\lambda}_{i+1}^T \cdot \left(\delta\underline{u}_{i+1} - \frac{\partial \underline{G}_i}{\partial \underline{u}_i} \cdot \delta\underline{u}_i - \frac{\partial \underline{G}_i}{\partial \underline{x}} \cdot \delta\underline{x} \right) \tag{A.10}$$

Rearranging all the terms leads to the following

$$\delta J = \left(\frac{\partial J}{\partial \underline{u}_n} + \underline{\lambda}_n^T \right) \cdot \delta\underline{u}_n + \sum_{i=1}^{n-1} \left(\underline{\lambda}_i^T - \underline{\lambda}_{i+1}^T \cdot \frac{\partial \underline{G}_i}{\partial \underline{u}_i} \right) \cdot \delta\underline{u}_i - \sum_{i=0}^{n-1} \underline{\lambda}_{i+1}^T \cdot \frac{\partial \underline{G}_i}{\partial \underline{x}} \cdot \delta\underline{x} \tag{A.11}$$

Recalling that our purpose is write the gradient of J as a function of only the control variables, it may be noticed that this goal is fulfilled if the multipliers $\underline{\lambda}_i$ are such that:

$$\begin{cases} \dfrac{\partial J}{\partial \underline{u}_n} + \underline{\lambda}_n^T = 0 \\ \underline{\lambda}_i^T - \underline{\lambda}_{i+1}^T \cdot \dfrac{\partial \underline{G}_i}{\partial \underline{u}_i} = 0 \qquad i = 1, \dots, n-1 \end{cases} \tag{A.12}$$

The equations A.12 are called adjoint equations, and, if they are verified, an infinitesimal variation of the objective function J may be expressed as:

$$\delta J = -\sum_{i=0}^{n-1} \underline{\lambda}_{i+1}^T \cdot \frac{\partial \underline{G}_i}{\partial \underline{x}} \cdot \delta\underline{x} \tag{A.13}$$

By the structure of the system of equations A.12 it can be noticed that the process is iterative, and, again, it is led from the backward, from the final state, determining $\underline{\lambda}_n$, to the initial state, which corresponds to $\underline{\lambda}_1$.

Adjoint method for the implicit functions

Let us consider an optimization problem of the form:

$$\begin{aligned} &min \qquad J = J(\underline{u}) \\ &s.t. \qquad \underline{G}(\underline{u}, \underline{x}) = 0 \end{aligned} \tag{A.14}$$

In general, for the state equations, it is difficult to find an explicit relation between $\underline{u}$ and $\underline{x}$ in order to operate a substitution in the expression of the objective function J. Nevertheless, keeping in mind the concept of Lagrange multipliers, it is possible to derive the gradient of J in function of the design variables $\underline{x}$ even in this case.

Again, let us linearize the objective function and the state equations:

$$\begin{aligned} &\delta J = \frac{\partial J}{\partial \underline{u}} \cdot \delta \underline{u} \\ &\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \cdot \delta \underline{u} + \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \delta \underline{x} = 0 \end{aligned} \tag{A.15}$$

In order to define an augmented expression of the small variation of J the vector of Lagrange multipliers $\underline{\lambda}$, so it is possible to write:

$$\begin{aligned} &\delta J = \frac{\partial J}{\partial \underline{u}} \cdot \delta \underline{u} + \underline{\lambda}^T \cdot \left(\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \cdot \delta \underline{u} + \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \delta \underline{x} \right) \quad \Rightarrow \\ &\delta J = \left(\frac{\partial J}{\partial \underline{u}} + \underline{\lambda}^T \cdot \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \right) \cdot \delta \underline{u} + \underline{\lambda}^T \cdot \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \delta \underline{x} \end{aligned} \tag{A.16}$$

Finally, in order to write the perturbation δJ as a function of only the infinitesimal variations of the design variables $\underline{x}$, it is necessary to solve the adjoint matrix equation:

$$\begin{aligned} &\frac{\partial J}{\partial \underline{u}} + \underline{\lambda}^T \cdot \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} = 0 \quad \Longrightarrow \quad \left(\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \right)^T \cdot \underline{\lambda} = -\left(\frac{\partial J}{\partial \underline{u}} \right)^T \quad \Longrightarrow \\ &\Longrightarrow \quad \underline{\lambda} = -\left(\left(\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \right)^T \right)^{-1} \cdot \left(\frac{\partial J}{\partial \underline{u}} \right)^T \end{aligned} \tag{A.17}$$

which, actually, represent a linear system in the variables $\underline{\lambda}$. If the Lagrange multipliers $\underline{\lambda}$ are calculated according to A.17, then the small perturbation of the objective function is:

$$\delta J = \underline{\lambda}^T \cdot \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \delta \underline{x} \tag{A.18}$$

Application of adjoint method to the optimization of the compliance in the case of a solid elastic continuum

Actually, adjoint method for the implicit functions is the one used in order to carry out the sensitivity analysis in the topology optimization of the elastic continuum structures, in the case of the minimization of the compliance. In fact, recalling the formulation of the topology optimization problem, and recalling that the state variables are the nodal displacements $\underline{u}$, and the control variables are the element densities $\underline{x} = \underline{\rho}$

$$\begin{aligned} &min \qquad J = \frac{1}{2} \cdot \left(\underline{u}(\underline{x})\right)^T \cdot \underline{\underline{K}}\left(\underline{\rho}\right) \cdot \underline{u}(\underline{x}) \\ &s.t. \qquad \underline{G}(\underline{u}, \underline{x}) = \underline{\underline{K}}(\underline{x}) \cdot \underline{u}(\underline{x}) - \underline{f} = 0 \end{aligned} \tag{A.19}$$

the adjoint equation is:

$$\underline{\lambda} = -\left(\left(\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{u}} \right)^T \right)^{-1} \cdot \left(\frac{\partial J}{\partial \underline{u}} \right)^T = -\left(\underline{\underline{K}}(\underline{x}) \right)^{-1} \cdot \frac{1}{2} \cdot 2 \cdot \underline{\underline{K}}(\underline{x}) \cdot \underline{u}(\underline{x}) = -\underline{u} \tag{A.20}$$

It may be notices that, actually, the Lagrange multipliers vector is equivalent to the vector of the state variables, despite the sign: for this reason, the optimization problem is said to be a self-adjoint problem. Furthermore, the derivative of the state equations with to respect the state variables is:

$$\frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} = \frac{\partial \underline{\underline{K}}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \underline{u} \tag{A.21}$$

Finally, the sensitivity of the objective function to the design variables, according to the equation A.18, is expressed by the following:

$$\delta J = \underline{\lambda}^T \cdot \frac{\partial \underline{G}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \delta \underline{x} = -\underline{\lambda}^T \cdot \frac{\partial \underline{\underline{K}}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \underline{u} \cdot \delta \underline{x} \; = -\underline{u}^T \cdot \frac{\partial \underline{\underline{K}}(\underline{u}, \underline{x})}{\partial \underline{x}} \cdot \underline{u} \cdot \delta \underline{x} \tag{A.22}$$

which, actually, is equivalent to the final equation 3.57.

Appendix B: equivalence between the mitigation of the volumetric error and the minimization of motor torque

It had been shown in chapter 4 that it is possible to optimize the performances of the redundant TRRTTT six axis machine tool using two different optimality criteria: the first one is the minimization of the position error due to the wrong angular position of the table motor. The correspondent objective function is expressed by the equation 4.43

$$U_1(H, \Omega, X, Z) = H^2 - 2Hr_P sin(\beta) + {r_P}^2$$

and the correspondent optimality condition imposed to the redundant axis reads like the equation 4.44

$$H_{opt1} = r_P sin(\beta)$$

The second optimality criterion is the minimization of the torque required to the motor of the turntable, when the tool applies a force $\underline{F}$ of components $F_x^{(6)}$ and $F_z^{(6)}$ on the workpiece. In this case, the optimality condition prescribes that the value of the axis H satisfies the equation 4.42:

$$\frac{F_z^{(6)}}{F_x^{(6)}} = \frac{\left(r_P \sin(\beta - \Omega) - H_{opt2} c\Omega\right)}{\left(r_P \cos(\beta - \Omega) - H_{opt2} s\Omega\right)}$$

Obviously, the value H_{opt2} depends by the inclination of the force, and, for this reason, it is not possible to define its value only in function of the position of the interaction point between tool and workpiece.

Nevertheless, it is possible to define an average index that, for every point P of the workspace, characterizes the effect of the position of the rotation axis of the table. Let us consider the configuration depicted in figure B.1: T is the projection of the rotation axis corresponding to a value H_T of the H axis. Furthermore, let us define the following index:

$$I = \int_0^{2\pi} F \cdot b \cdot |sin\theta| \cdot d\theta \qquad \text{(B.1)}$$

where F is a prescribed constant module of the force $\underline{F}$, applied to the point P, and b is the distance of the rotation axis from the line of action of $\underline{F}$ (moment arm).

Index I represents the integral of all the absolute value of the torques required to the motor when the slope θ of the force $\underline{F}$ covers all the range between 0 and 2π.

Integrating, it results:

$$I = \int_0^{2\pi} F \cdot b \cdot |sin\theta| \cdot d\theta = 2Fb \cdot \int_0^{\pi} sin\theta \cdot d\theta = \\ = 2Fb \cdot [-cos\theta]_{\theta=0}^{\theta=\pi} = 4Fb \tag{B.2}$$

Since the norm of the vector $\underline{F}$ is constant, the only way to minimize the average of all the contribution to the required motor torque, is minimizing the distance b. This means that the rotation axis must be in correspondence of the point O_1, ant this implies a condition equivalent to the optimality condition:

$$H_{opt} = H_{opt1} = r_P sin(\beta) = H_{opt1} \tag{B.3}$$

This demonstrates the substantial equivalence between the minimization of the volumetric error and the minimization of motor torque.

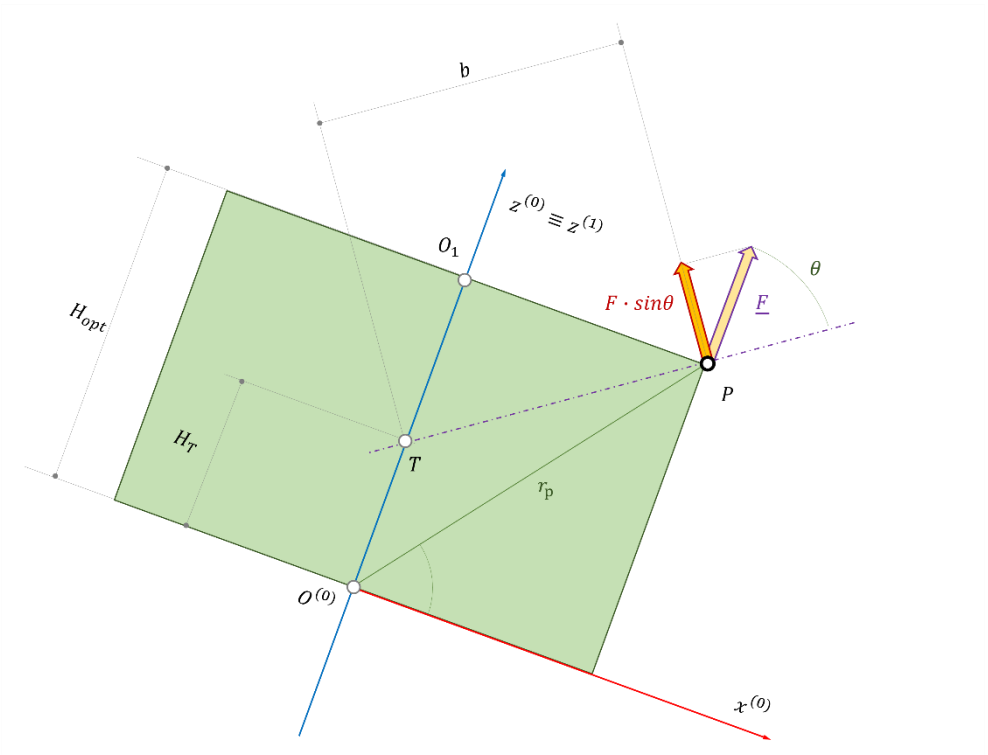


Figure B.1: generic force applied to a point of the workpiece by the tool

Appendix C: derivation of the single input / single output kinetic-static (KS) matrix

It will be derived now the relation between the input and output displacements, u_1 and u_2, and the correspondent input and output forces, f_1 and f_2. The required matrix equation should be of the following form:

$$\underline{\underline{H}} \cdot \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} = \begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix} \tag{C.1}$$

Two formulations will be provided below, the first one considering that the KS matrix $\underline{\underline{H}}$ can be expressed as a function of the mutual strain energy, and the mean compliance of the structure when subject to a dummy load at the input port, and a dummy load to the output port. This matrix equation should be of the following form:

$$\underline{\underline{H}} = \begin{vmatrix} H_1(k_{ij}) & H_s(k_{ij}) \\ H_s(k_{ij}) & H_2(k_{ij}) \end{vmatrix} \tag{C.2}$$

A second case, considers the matrix $\underline{\underline{H}}$ as a function of the element of the stiffness matrix $\underline{\underline{K}}$, reading as follows:

$$\underline{\underline{H}} = \begin{vmatrix} H_1(MC_1, MC_2, MSE) & H_s(MC_1, MC_2, MSE) \\ H_s(MC_1, MC_2, MSE) & H_2(MC_1, MC_2, MSE) \end{vmatrix} \tag{C.3}$$

KS matrix as function of the elements of the stiffness matrix

Regarding the first case, it is possible to make a reorder of the elements of the rigidity matrix $\underline{\underline{K}}$ as follows

$$\underline{\underline{K}} \cdot \underline{V_1} = \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \cdots & \cdots & \cdots \\ \vdots & & & \\ \vdots & & & \\ \vdots & & & \end{vmatrix} \cdot \begin{vmatrix} \begin{vmatrix} \delta_{11} \\ \delta_{21} \end{vmatrix} \\ \\ \\ \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \underline{F_1} \tag{C.4}$$

and

$$\underline{\underline{K}} \cdot \underline{V_2} = \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \cdots & \cdots & \cdots \\ \vdots & & & \\ \vdots & & & \\ \vdots & & & \end{vmatrix} \cdot \begin{vmatrix} \begin{vmatrix} \delta_{12} \\ \delta_{22} \end{vmatrix} \\ \\ \\ \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \underline{F_2} \tag{C.5}$$

If the material is linear elastic, and the displacement can be considered small, it is possible to apply the superposition of the effects, in order to obtain any combination of forces applied to the ports P_1 and P_2, being f_1 and f_2 the coefficient of the linear combination:

$$\underline{\underline{K}} \cdot \underline{V} = \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \cdots \\ \vdots & \end{vmatrix} \cdot \begin{vmatrix} \begin{vmatrix} f_1 \cdot \delta_{11} + f_1 \cdot \delta_{12} \\ f_2 \cdot \delta_{21} + f_2 \cdot \delta_{22} \end{vmatrix} \\ \begin{vmatrix} \vdots \end{vmatrix} \end{vmatrix} =$$

$$= \begin{vmatrix} \begin{vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{vmatrix} & \begin{vmatrix} \cdots & \underline{k}_{1s}{}^{T} & \cdots \end{vmatrix} \\ \begin{vmatrix} \vdots \\ \underline{k}_{1s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \vdots \\ \underline{k}_{2s} \\ \vdots \end{vmatrix} & \begin{vmatrix} \cdots & \underline{\underline{k}}_{ss} & \cdots \end{vmatrix} \end{vmatrix} \cdot \begin{vmatrix} \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} \\ \begin{vmatrix} \vdots \\ \underline{u}_s \\ \vdots \end{vmatrix} \end{vmatrix} = \begin{vmatrix} f_1 \\ f_2 \\ 0 \\ \vdots \\ 0 \end{vmatrix} = \underline{F} \tag{C.6}$$

Let us divide this matrix equation in two scalar equation and in a matrix equation:

$$\begin{cases} k_{11} \cdot u_1 + k_{12} \cdot u_2 + \underline{k}_{1s}{}^{T} \cdot \underline{u}_s = f_1 \\ k_{12} \cdot u_1 + k_{22} \cdot u_2 + \underline{k}_{2s}{}^{T} \cdot \underline{u}_s = f_2 \\ \underline{k}_{1s} \cdot u_1 + \underline{k}_{2s} \cdot u_2 + \underline{\underline{k}}_{ss} \cdot \underline{u}_s = \underline{0} \end{cases} \tag{C.7}$$

from the third equation:

$$\underline{k}_{1s} \cdot u_1 + \underline{k}_{2s} \cdot u_2 + \underline{\underline{k}}_{ss} \cdot \underline{u}_s = \underline{0} \quad \Rightarrow \quad \underline{u}_s = -\underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \cdot u_1 - \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \cdot u_2 \tag{C.8}$$

and making a substitution in the other two equations:

$$\begin{cases} k_{11} \cdot u_1 + k_{12} \cdot u_2 - \left(\underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \cdot u_1 + \underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \cdot u_2 \right) = f_1 \\ k_{12} \cdot u_1 + k_{22} \cdot u_2 - \left(\underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \cdot u_1 + \underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \cdot u_2 \right) = f_2 \\ \underline{u}_s = -\underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \cdot u_1 - \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \cdot u_2 \end{cases} \tag{C.9}$$

$$\Rightarrow$$

$$\begin{cases} \left(k_{11} - \underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \right) \cdot u_1 + \left(k_{12} - \underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \right) \cdot u_2 = f_1 \\ \left(k_{12} - \underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s} \right) \cdot u_1 + \left(k_{22} - \underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s} \right) \cdot u_2 = f_2 \end{cases}$$

This means that, is the structure is subject to a system of two forces (acting along two degrees of freedom), it is possible define a matrix equation involving only the two degrees of freedom [89]:

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix}$$

$where$:

$$\begin{cases} H_1 = \left(k_{11} - \underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s}\right) \\ H_2 = \left(k_{22} - \underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s}\right) \\ H_s = \left(k_{12} - \underline{k}_{1s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{2s}\right) = \left(k_{12} - \underline{k}_{2s}{}^{T} \cdot \underline{\underline{k}}_{ss}{}^{-1} \cdot \underline{k}_{1s}\right) \end{cases} \quad \text{(C.10)}$$

KS matrix as function of input and output mean compliances and mutual strain energy

In order to formulate the $\underline{\underline{H}}$ matrix as an expression of MC_1, MC_2, and MSE, let us recall that

$$\begin{cases} MC_1 = f_1 \cdot \delta_{11} = \delta_{11} = \underline{V_1}{}^{T} \cdot \underline{\underline{K}} \cdot \underline{V_1} \\ MC_2 = f_2 \cdot \delta_{22} = \delta_{22} = \underline{V_2}{}^{T} \cdot \underline{\underline{K}} \cdot \underline{V_2} \\ MSE = MSE_1 = \underline{V_1}{}^{T} \cdot \underline{\underline{K}} \cdot \underline{V_2} = \delta_{12} = MSE_2 = \underline{V_2}{}^{T} \cdot \underline{\underline{K}} \cdot \underline{V_1} = \delta_{21} \end{cases} \quad \text{(C.11)}$$

and applying the new formulation of the stiffness matrix in the two cases of structure subject to one dummy load at a time:

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} \delta_{11} \\ \delta_{12} \end{vmatrix} = \begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} MC_1 \\ MSE \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \end{vmatrix}$$

and (C.12)

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} \delta_{21} \\ \delta_{22} \end{vmatrix} = \begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} MSE \\ MC_2 \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \end{vmatrix}$$

It is possible to write the coefficients R_1, R_1, and R_1 in function of SE_1, SE_2, and MSE:

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} MC_1 \\ MSE \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \end{vmatrix} \Rightarrow \begin{cases} H_1 \cdot MC_1 + H_s \cdot MSE = 1 \\ H_s \cdot MC_1 + H_2 \cdot MSE = 0 \end{cases} \Rightarrow \quad H_s = -\frac{MSE}{MC_1} \cdot H_2 \Rightarrow \quad \text{(C.13)}$$

$$\Rightarrow \quad H_1 \cdot MC_1 - \frac{MSE}{MC_1} \cdot H_2 \cdot MSE = MC_1 \cdot H_1 - \frac{MSE^2}{MC_1} \cdot H_2 = 1$$

and

$$\begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} \cdot \begin{vmatrix} MSE \\ MC_2 \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \end{vmatrix} \Rightarrow \begin{cases} H_1 \cdot MSE + H_s \cdot MC_2 = 0 \Rightarrow \\ H_s \cdot MSE + H_2 \cdot MC_2 = 1 \end{cases} \quad H_s = -\frac{MSE}{MC_2} \cdot H_1 \Rightarrow$$
$$\Rightarrow \quad -\frac{MSE}{MC_2} \cdot H_1 \cdot MSE + H_2 \cdot MC_2 = -\frac{MSE^2}{MC_2} \cdot H_1 + MC_2 \cdot H_2 = 1 \tag{C.14}$$

Solving with to respect R_1 and R_2:

$$\begin{cases} MC_1 \cdot H_1 - \frac{MSE^2}{MC_1} \cdot H_2 = 1 \\ -\frac{MSE^2}{MC_2} \cdot H_1 + MC_2 \cdot H_2 = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \frac{MC_1 \cdot MC_2}{MSE^2} \cdot MC_1 \cdot H_1 - \frac{MC_1 \cdot MC_2}{MSE^2} \cdot \frac{MSE^2}{MC_1} \cdot H_2 = \frac{MC_1 \cdot MC_2}{MSE^2} \\ -\frac{MSE^2}{MC_2} \cdot H_1 + MC_2 \cdot H_2 = 1 \end{cases} \Rightarrow \tag{C.15}$$

$$\Rightarrow \quad \frac{MC_1 \cdot MC_2}{MSE^2} \cdot MC_1 \cdot H_1 - \frac{MSE^2}{MC_2} \cdot H_1 = \frac{MC_1 \cdot MC_2}{MSE^2} + 1 \quad \Rightarrow$$

$$\Rightarrow \quad H_1 = \frac{MC_2}{MC_1 \cdot MC_2 - MSE^2}$$

and

$$\begin{cases} MC_1 \cdot H_1 - \frac{MSE^2}{MC_1} \cdot H_2 = 1 \\ -\frac{MSE^2}{MC_2} \cdot H_1 + MC_2 \cdot H_2 = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} MC_1 \cdot H_1 - \frac{MSE^2}{MC_1} \cdot H_2 = 1 \\ -\frac{MC_1 \cdot MC_2}{MSE^2} \cdot \frac{MSE^2}{MC_2} \cdot H_1 + \frac{MC_1 \cdot MC_2}{MSE^2} \cdot MC_2 \cdot H_2 = \frac{MC_1 \cdot MC_2}{MSE^2} \end{cases} \Rightarrow \tag{C.16}$$

$$\Rightarrow \quad -\frac{MSE^2}{MC_1} \cdot H_2 + \frac{MC_1 \cdot MC_2}{MSE^2} \cdot MC_2 \cdot H_2 = \frac{MC_1 \cdot MC_2}{MSE^2} + 1 \quad \Rightarrow$$

$$\Rightarrow \quad H_2 = \frac{MC_1}{MC_1 \cdot MC_2 - MSE^2}$$

Furthermore,

$$H_s = -\frac{MSE}{MC_2} \cdot H_1 = -\frac{MSE}{MC_1} \cdot H_2 = -\frac{MSE}{MC_1 \cdot MC_2 - MSE^2} \tag{C.17}$$

Finally, it is possible to write the matrix $\underline{\underline{R}}$:

$$\underline{\underline{R}} = \begin{vmatrix} H_1 & H_s \\ H_s & H_2 \end{vmatrix} = \frac{1}{MC_1 \cdot MC_2 - MSE^2} \cdot \begin{vmatrix} SE_2 & -MSE \\ -MSE & SE_1 \end{vmatrix} \tag{C.18}$$

so that:

$$\frac{1}{MC_1 \cdot MC_2 - MSE^2} \cdot \begin{vmatrix} MC_2 & -MSE \\ -MSE & MC_1 \end{vmatrix} \cdot \begin{vmatrix} u_1 \\ u_2 \end{vmatrix} = \begin{vmatrix} f_1 \\ f_2 \end{vmatrix} \tag{C.19}$$

for a generic structure subject to two single components forces.

Appendix D: resume table of the objective function for the synthesis of compliant mechanisms by the mean of continuum topology optimization

	Objective Function	Model	Approach	Mathematical Programming Algorithm	Hinges avoidance	Reference(s)
(1)	$-\alpha MSE + (1-\alpha)SE$					[84]
(2)	$GA = \frac{MSE}{SE}$	Output Spring	Level set	Gradient Projection	No	[84][89]
(3)	MA	Output Spring	Level set, ESO	No MP Algorithm (Rejection criterium)	-	[89][91]
(4)	$WR \quad (\equiv MA \cdot GA)$	Output Spring	SIMP	MMA	No	[86]
(5)	$\frac{GA}{SE}$	Output Spring	BESO	No MP Algorithm (Bisection algorithm)	Yes	[94]

(6)	$(WR \equiv)\ \ MA \cdot GA$	Input Spring and Output Spring	BESO	No MP Algorithm (Bisection algorithm)	-	[92]g
(7)	Δ_{out}	Input Spring and Output Spring	BESO	No MP Algorithm (Bisection algorithm)	Yes	[93]
(8)	$-u_{OUT} + \lambda(C^* + C)$	Output Spring	BESO	No MP Algorithm (Bisection algorithm)	No	[95]
(9)	$\frac{MSE}{(SE_1 + SE_2)}$	No Spring: characteristic stiffness	Homogenization Method	Sequential Linear Programming	-	[88]
(10)	$-u_{OUT} + \alpha C_{IN} + \beta C_{OUT}$	No Spring: characteristic stiffness	SIMP	Optimality Criteria	Yes	[96]
(11)	$e^{-(GA-GA^*)} \cdot K_i \cdot K_o$	No Spring: characteristic stiffness	Level Set	Gradient descent	Yes	[99]
(12)	$\sum \varepsilon_i$		SIMP	(Physical Programming Method)	Yes	[97]
(13)	$\frac{MSE}{\Phi}$	No Spring: distributed deformation	SIMP	Sequential Linear Programming	Yes	[98]