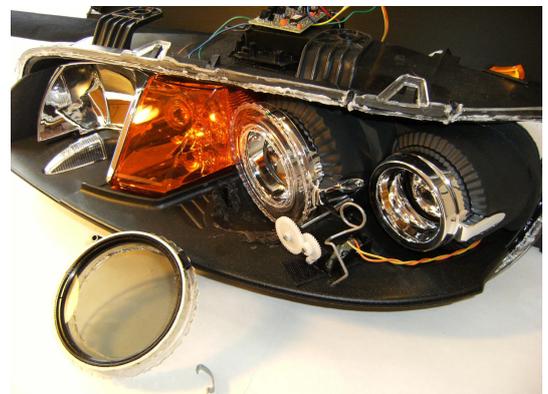
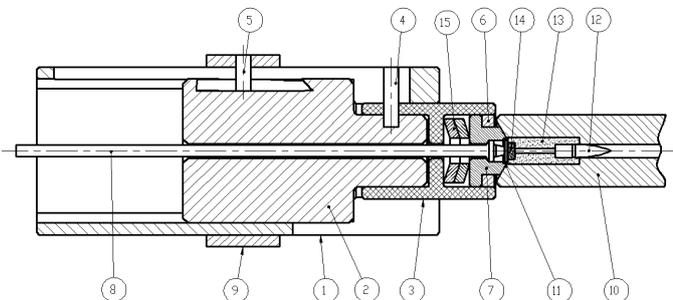
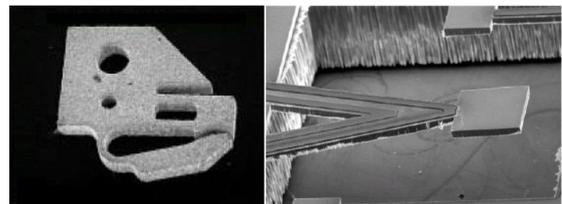
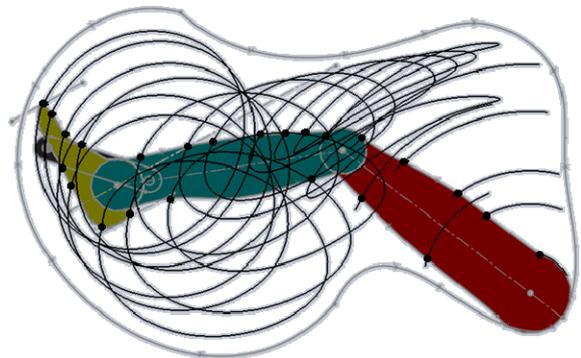


International Journal of Mechanics and Control

Editor: Andrea Manuello Bertetto



Editorial Board of the
International Journal of Mechanics and Control

Published by Levrotto&Bella – Torino – Italy E.C.

Honorary editors

Guido Belforte

Kazy Yamafuji

Editor: Andrea Manuello Bertetto

General Secretariat: Elvio Bonisoli

Atlas Akhmetzyanov
*V.A.Trapeznikov Institute of Control Sciences
of Russian Academy of Sciences
Moscow – Russia*

Domenico Appendino
*Prima Industrie
Torino – Italy*

Kenji Araki
*Saitama University
Shimo Okubo, Urawa
Saitama – Japan*

Guido Belforte
*Technical University – Politecnico di Torino
Torino – Italy*

Bruno A. Boley
*Columbia University,
New York – USA*

Marco Ceccarelli
*LARM at DIMSAT
University of Cassino
Cassino – Italy*

Amalia Ercoli Finzi
*Technical University – Politecnico di Milano
Milano – Italy*

Carlo Ferraresi
*Technical University – Politecnico di Torino
Torino – Italy*

Anindya Ghoshal
*Arizona State University
Tempe – Arizona – USA*

Nunziatino Gualtieri
*Space System Group
Alenia Spazio
Torino – Italy*

Alexandre Ivanov
*Technical University – Politecnico di Torino
Torino – Italy*

Giovanni Jacazio
*Technical University – Politecnico di Torino
Torino – Italy*

Takashi Kawamura
*Shinshu University
Nagano – Japan*

Kin Huat Low
*School of Mechanical and Aerospace Engineering
Nanyang Technological University
Singapore*

Andrea Manuello Bertetto
*University of Cagliari
Cagliari – Italy*

Stamos Papastergiou
*Jet Joint Undertaking
Abingdon – United Kingdom*

Mihailo Ristic
*Imperial College
London – United Kingdom*

János Somló
*Technical University of Budapest
Budapest – Hungary*

Jozef Suchy
*Faculty of Natural Science
Banska Bystrica – Slovakia*

Federico Thomas
*Instituto de Robótica e Informática Industrial
(CSIC-UPC)
Barcelona – Espana*

Lubomir Uher
*Institute of Control Theory and Robotics
Bratislava – Slovakia*

Furio Vatta
*Technical University – Politecnico di Torino
Torino – Italy*

Vladimir Viktorov
*Technical University – Politecnico di Torino
Torino – Italy*

Kazy Yamafuji
*University of Electro-Communications
Tokyo – Japan*

*Official Torino Italy Court Registration
n.5390, 5th May 2000*

*Deposito presso il Tribunale di Torino
numero 5390 del 5 maggio 2000*

Direttore responsabile:

Andrea Manuello Bertetto

International Journal of Mechanics and Control

Editor: Andrea Manuello Bertetto

***Honorary editors: Guido Belforte
Kazy Yamafuji***

General Secretariat: Elvio Bonisoli

The Journal is addressed to scientists and engineers who work in the fields of mechanics (mechanics, machines, systems, control, structures). It is edited in Turin (Northern Italy) by Levrotto&Bella Co., with an international board of editors. It will have not advertising.

Turin has a great and long tradition in mechanics and automation of mechanical systems. The journal would will to satisfy the needs of young research workers of having their work published on a qualified paper in a short time, and of the public need to read the results of researches as fast as possible.

Interested parties will be University Departments, Private or Public Research Centres, Innovative Industries.

Aims and scope

The *International Journal of Mechanics and Control* publishes as rapidly as possible manuscripts of high standards. It aims at providing a fast means of exchange of ideas among workers in Mechanics, at offering an effective method of bringing new results quickly to the public and at establishing an informal vehicle for the discussion of ideas that may still in the formative stages.

Language: English

International Journal of Mechanics and Control will publish both scientific and applied contributions. The scope of the journal includes theoretical and computational methods, their applications and experimental procedures used to validate the theoretical foundations. The research reported in the journal will address the issues of new formulations, solution, algorithms, computational efficiency, analytical and computational kinematics synthesis, system dynamics, structures, flexibility effects, control, optimisation, real-time simulation, reliability and durability. Fields such as vehicle dynamics, aerospace technology, robotics and mechatronics, machine dynamics, crashworthiness, biomechanics, computer graphics, or system identification are also covered by the journal.

Please address contributions to

Prof. Guido Belforte
Prof. Andrea Manuello Bertetto
PhD Eng. Elvio Bonisoli

*Dept. of Mechanics
Technical University - Politecnico di Torino
C.so Duca degli Abruzzi, 24.
10129 - Torino - Italy - E.C.*

www.jomac.it
e_mail: jomac@polito.it

Subscription information

Subscription order must be sent to the publisher:

*Libreria Editrice Universitaria
Levrotto&Bella
2/E via Pigafetta – 10129 Torino – Italy*

www.levrotto-bella.net
e_mail: info@levrotto-bella.net
tel. +39.011.5097367
+39.011.5083690
fax +39.011.504025

MODELLING AND TEMPERATURE CONTROL OF SHAPE MEMORY ALLOYS WITH FAST ELECTRICAL HEATING

Ramiro Velázquez^{1,2} Edwige E. Pissaloux²

¹ Mecatrónica y Control de Sistemas, Universidad Panamericana

² Institut des Systèmes Intelligents et de Robotique (ISIR), CNRS UMR 7222 - Université Paris 6

ABSTRACT

This paper presents a comparative theoretical study of the performance of a set of controllers for improving the speed of actuators based on shape memory alloys (SMAs), especially on Nitinol (NiTi). To prevent overheating and thermal fatigue, these controllers take into account the maximum heating current at which a NiTi element can safely be heated. The thermal behaviour of NiTi is first modelled for calculating the time response and, based on the suggested model, it is shown that hysteresis due to phase transformation can be neglected for rapid heating, thus simplifying the model to a linear problem. The design and performance of a set of linear controllers is then presented. Simulation results show a substantial increase in actuation speed.

Keywords: Actuation speed, modelling, NiTi wire, shape memory alloys (SMAs), temperature control.

1 INTRODUCTION

The interest on smart materials such as shape memory alloys (SMAs), piezoelectric ceramics, electroactive polymers (EAPs) and electrorheological (ER) fluids has grown in the last decades due to their unique properties. This class of materials has the ability of changing its shape, size, stiffness, among other properties, through the imposition of electrical, electromagnetic, temperature or stress fields.

In particular, SMAs are metallic alloys that are able to recover their original shape when subjected to the appropriate thermal process: at low temperature martensite (*M*) phase, the crystalline structure of a SMA is quite malleable and can be easily deformed into new arrangements, which are retained. Upon heating, the crystalline structure returns, in a high temperature austenite (*A*) phase, to a “memorized” configuration (Fig. 1).

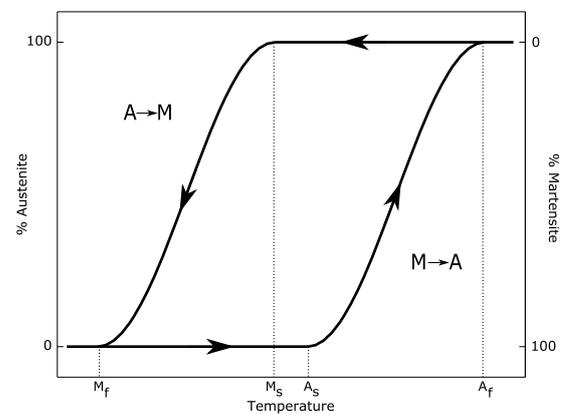


Figure 1 Schematic of M→A and A→M transformations, after [1].

The M→A and A→M transformations are essentially characterized by four temperatures: A_s and A_f are the start and finish temperatures for the *A* phase, respectively, while M_s and M_f are the start and finish temperatures for the *M* phase, respectively. As seen, SMAs show a significant hysteresis during both transformations.

Using SMAs provides an interesting actuation principle: compact size, high power/ weight ratio, extremely high fatigue resistance to cyclic operation and smooth, clean, spark-free and noiseless performance [2]. These remarkable properties have been attracting growing interest in several fields of sciences and engineering.

Contact author: Ramiro Velázquez¹, Edwige E. Pissaloux²

¹ Fracc. Rústicos Calpulli. C.P. 20290, Aguascalientes, Mexico, E-mail: rvelazquez@ags.up.mx

² Pyramide tour 55, 4 place Jussieu, 75252 Paris Cedex 05, France, E-mail: edwige.pissaloux@upmc.fr

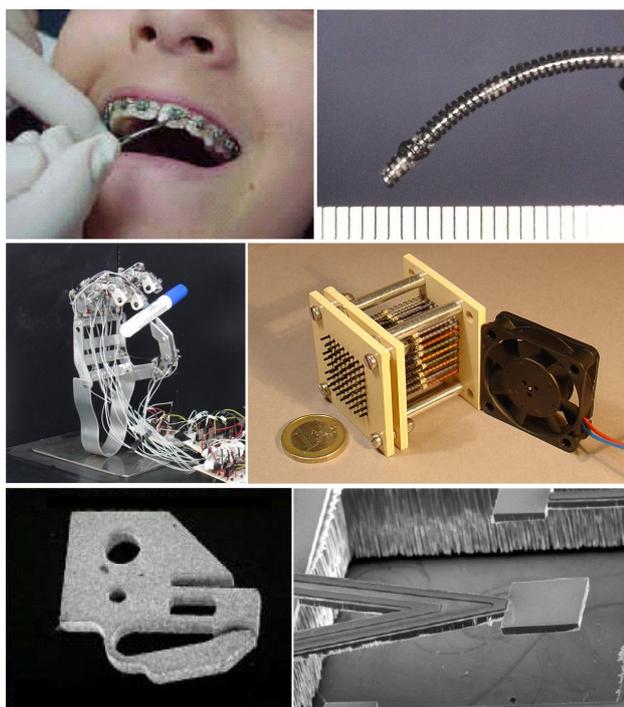


Figure 2 SMA-based devices - (top) Biomedical field: orthodontic wire and endoscope for intestinal exploration. (middle) Robotics field: four-fingered robot hand and portable tactile display for the blind. (bottom) MEMS field: micro-gripper and micro-mirror.

SMA's have become successful for biomedical applications due to their excellent bio-compatibility and non-invasive properties. They have been employed in orthodontic devices [3], surgical instruments such as endoscopes [4], among other applications (Fig. 2 top).

Robotics is another popular application field for SMA's. Basically, they are used as actuators that expand and stretch creating linear movement. Robotic hands [5], tactile displays [6], among other devices have been developed using SMA's (Fig. 2 middle).

SMA's are also used as actuators in micro-electromechanical systems (MEMS). Micro-grippers [7] and micro-mirrors [8] are some examples of SMA-based MEMS (Fig. 2 bottom).

Although their attractive properties, SMA's have two major drawbacks: their relatively slow response speed and their non-linear behaviour due to the hysteresis shown in fig. 1.

The limiting factors on the speed of SMA's are the heating and cooling rates. The cooling rate can be increased by various means such as forced-air convection, oil or water cooling. For heating, practically any source of heat producing a temperature change in the material and/or its surrounding medium can be used: electrical current, Peltier effect, laser, hot air, hot fluid, etc. When electrically driven, the heating rate can only be increased by passing a larger current through the SMA.

In principle, an SMA can be heated arbitrarily quickly by passing a sufficiently large current through it. However,

excessive electrical power has the capacity to overheat the SMA, causing thermal stress fatigue and a gradual degradation of its performance.

Some micro-structural observations performed in a previous work [9] show that SMA's are in fact very sensitive to overheating. Fig. 3 compares the micro-structures of a Nitinol (NiTi) SMA before and after thermal fatigue. Fig. 3(a) shows the surface of a brand new NiTi SMA wire. Note that the material presents a particular homogeneity in its structure. Fig. 3(b) shows the surface of the same specimen right after 20 overheating cycles. Note that overheating actually induces micro-structural modifications: a new granular order is observed in SMA's. Fig. 3(c) shows the surface of the same specimen after 50 overheating cycles. Note that the granular distribution becomes more accentuated with bigger and more defined grains. As suggested in [10], SMA's exhibit a strong dependence between grain size and performance: as the grain size increases, the strength of the material decreases.

To avoid thermal fatigue, a strategy for rapid electrical heating that avoids overheating must be adopted. This paper presents a simple analytical model to predict the SMA thermal behaviour and compares the performance of a set of controllers designed to increase actuation speed and avoid overheating.

The rest of the paper is organized as follows: Section 2 introduces the SMA thermal dynamics and discusses the main parameters characterization. Section 3 presents the design and performance of several temperature controllers while Section 4 discusses some practical considerations for these controllers. Finally, Section 5 concludes the paper summarizing main concepts and results.

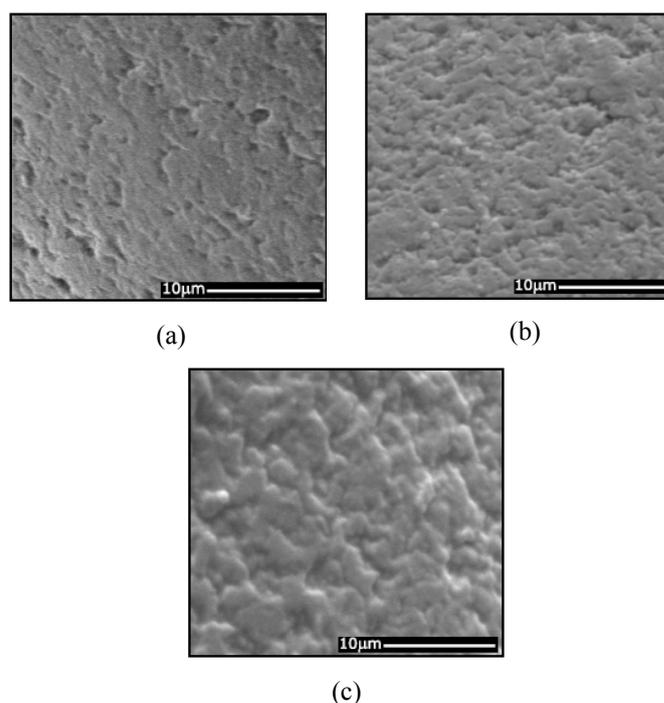


Figure 3 Comparison between the micro-structures of NiTi SMA's before and after thermal fatigue due to overheating.

2 MODELLING THE THERMAL DYNAMICS OF SMAs

Extensive work has been devoted to model the SMA behaviour. Approaches range from atomic interactions [11], [12] and thermodynamic formulations [13], [14] to phenomenological models based on experimental data [15], [16]. Models agree, regardless of their approach, that a SMA element can be considered as a three-element system in which thermal energy is converted into mechanical work (Fig. 4).

In spite of their high quality, most of these models are difficult to use in practice: they lead to quite complex mathematical equations that require burdensome numerical solution. Moreover, they involve parameters not clearly identified and depend on previously determined experimental data, which results totally impractical for rapid evaluation or design.

This section discusses the main features associated with SMA thermal modelling through existing constitutive models suitable for easy analytical prediction of the SMA behaviour: they use only the geometry of the SMA element and few material properties extracted from the manufacturer's data sheets or from the literature.

2.1 THERMAL MODEL

Let us start analyzing the SMA thermal behaviour considering a system in thermal equilibrium where Q_s is the heat generated by an external source and $Q_{conduction}$, $Q_{convection}$ and $Q_{radiation}$ are the different dissipation forms of heat. From this thermal balance, the heat stored in the SMA can be expressed using the classical differential equation of heat transfer as follows:

$$\rho CV \frac{dT}{dt} = Q_s - Q_{conduction} - Q_{convection} - Q_{radiation} \quad (1)$$

where ρ is the density of the SMA, C is its specific heat, V is the volume of the SMA material and T is the temperature of the SMA at a time t .

For electrically driven SMAs, heat is generated in accordance with the Lenz-Joule law:

$$Q_s = i^2 R \quad (2)$$

where i is the electrical current applied and R is the element's electrical resistance. For most applications (those within the atmosphere), $Q_{radiation}$ can be neglected. Concerning $Q_{convection}$, it can be expressed as:

$$Q_{convection} = hs(T - T_e) \quad (3)$$

where h is the heat-exchange coefficient between the SMA and the surrounding medium, S is the surface area of the SMA and T_e is the environment's temperature.

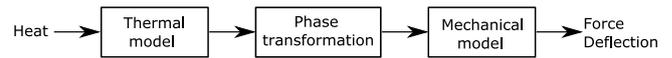


Figure 4 SMA block diagram model.

Being a metallic alloy, $Q_{conduction}$ is usually neglected by assuming that the temperature is uniform in the material at all times due to its low internal resistance to heat conduction. A fast method to evaluate when $Q_{conduction}$ can be neglected is the Biot number defined by:

$$B_i = \frac{hl}{k} \quad (4)$$

where l is the characteristic length of the SMA material and k is its thermal conductivity. If B_i is less than 0.01, then this assumption is considered valid.

Once the Biot number has been validated, eq. (1) becomes:

$$\rho CV \frac{dT}{dt} = i^2 R - hs(T - T_e) \quad (5)$$

The block diagram model of the SMA thermal behaviour is shown in fig. 5.

2.2 HEAT-EXCHANGE COEFFICIENT

Recall that, in classical theory of convective heat transfer, the heat-exchange coefficient h is defined by:

$$h = \frac{\lambda N_u}{l} \quad (6)$$

where λ is the thermal conductivity of the convective medium (gas or fluid) and N_u is the Nusselt number which defines the heat-exchange ratio given the conditions of convection (free or forced). For free convection, N_u is typically expressed as [17]:

$$N_u = \left[0.6 + \frac{0.387(G_r P_r)^{\frac{1}{6}}}{\left[1 + (0.559/P_r)^{\frac{9}{16}} \right]^{\frac{8}{27}}} \right]^2 \quad (7)$$

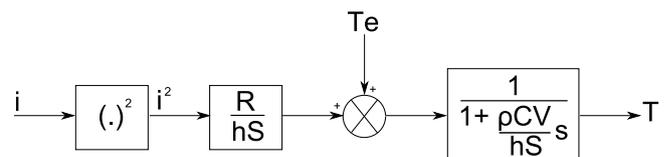


Figure 5 Functional block diagram of the SMA thermal behaviour.

where P_r and G_r are the Prandtl and Grashof numbers for the convective medium, respectively.

Concerning the characteristic length l , it is commonly defined as the volume of the material divided by its surface area ($l=V/S$). However, in the case of thin metallic cylinders or wires (a particularly popular shape in SMAs), it is more pertinent to consider the characteristic length equal to the cylinder's diameter ($l=d$).

2.3 SIMULATION RESULTS

As illustrative example, consider a NiTi SMA wire of 200 μm diameter and 7 mm length. The thermal properties of NiTi, directly extracted from the manufacturer's data sheets [18], as well as the air properties of interest for this study are given in table I. The rest of the parameters needed for simulation are also shown in table I. They can either be calculated from those in table I or easily measured through simple experiments. Being the Biot number valid, eq. (5) can be used.

Fig. 6 shows the time response of the NiTi wire to a step current of 300 mA. Recall that, once temperature achieves A_f , the transformation is complete and the SMA has recovered its memorized shape. Any additional heat supplied to the SMA is useless and overheats the material. Furthermore, the SMA's response becomes problematic: as frequency increases, the cooling rate decreases significantly enough so that the A \rightarrow M transformation cannot be completed. The SMA goes into saturation and the memory effect is no longer observed.

Fig. 7 shows the time response of the NiTi wire to a 1 Hz, 300 mA square signal. Note that M_f is never achieved.

Table I - Geometric/thermal properties of NiTi and surrounding medium.

Property	Value	Unit
NiTi properties		
Density ρ	6450	$\text{kg}\cdot\text{m}^{-3}$
Specific heat C	320	$\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$
Thermal conductivity k	8	$\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
Transformation temperatures A_f, A_s, M_s, M_f	78, 68, 52, 42	$^{\circ}\text{C}$
Air properties (20$^{\circ}\text{C}$)		
Thermal conductivity λ	0.0257	$\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$
Grashof number G_r	0.3903	-
Prandtl number P_r	0.713	-
Calculated/measured properties		
Surface S	4.46×10^{-6}	m^2
Volume V	2.2×10^{-10}	m^3
Nusselt number N_u	0.739	-
Characteristic length l	200×10^{-6}	m
Heat-exchange coefficient h	94.96	$\text{W}\cdot\text{m}^{-2}\cdot\text{K}^{-1}$
Resistance R	3	Ω
Environment's temperature T_e	20	$^{\circ}\text{C}$

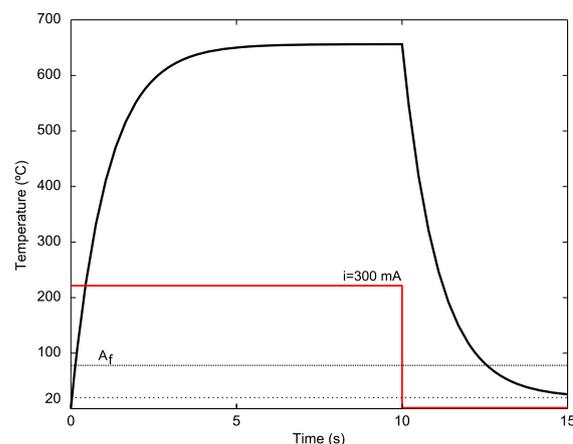


Figure 6 Temperature-time response to a step input current of 300 mA.

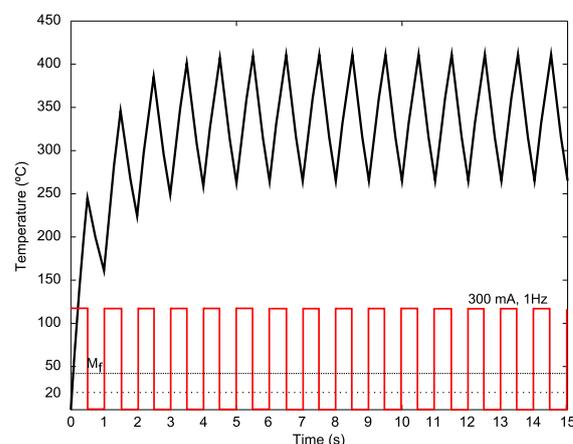


Figure 7 Response to a 1 Hz, 300 mA square signal.

2.4 PHASE TRANSFORMATION

Unfortunately, the actual SMA thermal behaviour is more complex. As seen in fig. 1, SMAs, unlike traditional materials, show hysteresis during both M \rightarrow A and A \rightarrow M transformations. Physically, these hystereses are a dissipation and assimilation of latent heat due to phase transformation that tend to slow down both heating and cooling processes. Consequently, it is pertinent to evaluate this effect on the thermal dynamics.

A successful empirical relation has been proposed by Liang and Rogers [1], which describes the amount of martensite fraction ξ transformed on a temperature T . For constant load conditions, ξ can be written for heating as:

$$\xi(T) = \begin{cases} 0 & \text{at } T > A_f \\ \frac{1}{2} \left[\cos \left(\pi \frac{T - A_s}{A_f - A_s} \right) + 1 \right] & \text{at } A_s < T < A_f \\ 1 & \text{at } T < A_s \end{cases} \quad (8)$$

The analogous dependence can be expressed for cooling by replacing A_s and A_f temperatures with M_f and M_s temperatures correspondingly.

Eq. (8) completes the energy balance eq. (5) with the phase transformation parameter as follows:

$$\rho V \left[C \frac{dT}{dt} + \Delta H \frac{d\xi}{dt} \right] = i^2 R - hS(T - T_e) \quad (9)$$

where ΔH is the latent heat defined on the transformation interval. Solution of eq. (9) can be obtained using any scientific computing software such as *Maple* (Maplesoft Co.) or *Mathematica* (Wolfram Research). For example, the analytical solution for heating in $A_s < T < A_f$ obtained using *Mathematica* can be expressed as eq. (10):

$$t = t_1 + \frac{\rho CV}{hS} \left[\ln \left(\frac{-hS}{V} A_s + Q \right) - \frac{\Delta H}{2} e^U E_i(1, U) - C \ln \left(\frac{-hS}{V} A_f + Q \right) + \frac{\Delta H}{2} e^J E_i(1, J) \right]$$

with:

$$\begin{aligned} Q &= \frac{hS}{V} T_e + i^2 R \\ U &= \frac{1}{2} (A_s - T_e - \frac{V}{hS} i^2 R) \\ J &= \frac{1}{2} (A_f - T_e - \frac{V}{hS} i^2 R) \\ E_i(n, x) &= \int_1^{\infty} \frac{e^{-xt}}{t^n} dt \end{aligned} \quad (10)$$

where t_1 is the time to reach A_s . Similarly, the analytical solution for cooling ($i^2 R = 0$) in $M_f < T < M_s$ can be expressed as eq. (11):

$$t = t_2 + \frac{\pi \rho V \Delta H}{2 hS \Psi} \left\{ \cos(\pi X) [Si(\pi Y) - Si(\pi Z)] - \sin(\pi X) [Ci(\pi Y) - Ci(\pi Z)] \right\}$$

with:

$$\begin{aligned} X &= \frac{M_f - T_e}{\Psi} \\ Y &= \frac{T - T_e}{\Psi} \\ Z &= \frac{M_s - T_e}{\Psi} \\ \Psi &= M_s - M_f \\ Si(x) &= \int_0^x \frac{\sin(y)}{y} dy \\ Ci(x) &= \int_0^x \frac{\cos(y)}{y} dy \end{aligned} \quad (11)$$

where t_2 is the time to reach M_s (i.e. before transformation starts).

Simplified analytical methods for solving eq. (9) do exist in the literature [19], [20] and might be more practical, faster and easier to implement.

Fig. 8 shows the delay effect due to phase transformation on the thermal response of the 200 μm diameter wire to a step current of 300 mA. The latent heat of transformation ΔH was assumed to be 14 J/g, a typical value in NiTi SMAs. Note that the effect of phase transformation moderately extends the heating time, but significantly increases the cooling one.

Fig. 9 shows the time response to a set of step currents applied to the 200 μm NiTi wire. The amplitude is varied from 200 to 800 mA. Note that the greater the current, the less the influence of the latent heat on the actuation time. It can be concluded that the delay effect due to phase transformation can be reasonably neglected especially for input current intensities higher than 800 mA. Thus, the SMA thermal behaviour can be simplified to a linear model and the functional block diagram of fig. 5 can be retained.

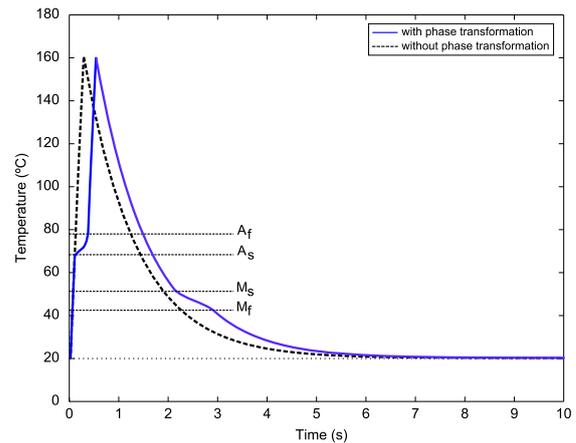


Figure 8 Comparison of temperature-time curves for the NiTi wire: with and without phase transformation.

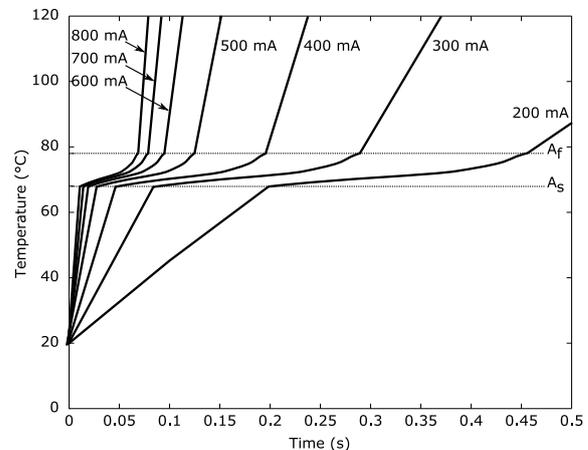


Figure 9 M \rightarrow A phase transformation of the NiTi wire at different heating electric currents (200 to 800 mA).

3 TEMPERATURE CONTROL

As aforementioned, the only way to increase the heating rate of an electrically driven SMA is by passing larger currents through it. However, the risk of overheating, damaging and saturation does exist. A closed-loop control scheme is necessary to regulate the thermal evolution of the SMA.

A temperature controller for SMAs should accomplish the two following preliminary criteria:

- Make the SMA achieve the A_f temperature precisely and as fast as possible.
- Present no overshoot. An oscillatory response should be avoided because it implies cooling. No control on the cooling process can be designed or performed as it only depends on air convection.

Four linear controllers have been tested for temperature control of the NiTi wire: Proportional (P), Proportional-Integral (PI), Proportional-Integral-Derivative (PID) and Feed-forward control.

For the first three controllers, gains were adjusted to meet the criteria established above while the control scheme in fig. 10 was used for feed-forward control.

Fig. 12(a) compares all four responses while table II summarizes the main performance specifications. Note that feed-forward control offers the best performance: the fastest and most precise response.

Fig. 12(b) compares the driving input currents generated by each controller for achieving the responses of fig. 12(a). Note that fast actuation requires large currents: the P-controller requires 1.6 A, PI 3.2 A, PID 16 A and feed-forward control requires up to 24 A.

4 PRACTICAL CONSIDERATIONS

Unfortunately, the temperature-time responses in fig. 12(a) cannot be achieved in practice because of the current magnitudes of fig. 12(b): even though these currents are applied just for a very short time, they are large enough to damage the NiTi wire.

One possible solution to this problem is to limit the current supplied from the controller to the SMA. Fig. 11 shows one possible way to incorporate the rapid heating method into a temperature control system while keeping the current safe for the SMA [21].

Suitable values for i_{safe} can be found in data sheets for NiTi wires or determined by simple experimentation. Another important parameter that can define the value of i_{safe} is the available current from the power supply.

Fig. 12(c) compares the step responses of all four controllers with a current limiter. The maximum current value allowed to the NiTi wire is 1 A. Table III summarizes the main performance specifications for this case. Note that the performances of the P and PI controllers are practically identical to those in table II. Note also, that the PID controller offers now the best performance in precision and actuation speed.

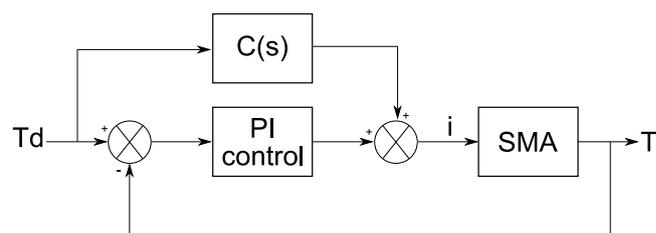


Figure 10 Feed-forward control scheme. A PI controller was used for regulating the main loop. $C(s)$ is a lead/lag function which regulates the feed-forward loop.

Table II - Summary of specifications and performances.

	Controller	Settling time	Steady error
P	$k_p=0.02$	0.15 s	1.75%
PI	$k_p=0.04$ $k_i=0.037$	0.08 s	0.22%
PID	$k_p=0.2$ $k_d=2 \times 10^{-10}$ $k_i=1.07$	0.021 s	1%
Feed-forward	$C(s)=\frac{10^{-4}(0.26s+0.5)}{1.07s+1}$	0.014 s	0.01%

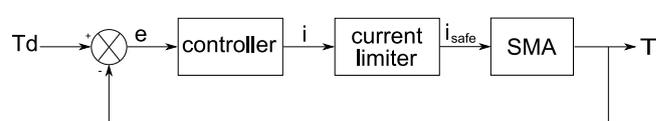


Figure 11 Temperature control system with current limiter.

Table III - Summary of performances with current limiter block.

	Settling time	Steady error
P	0.15 s	1.75%
PI	0.08 s	1.06%
PID	0.032 s	0.4%
Feed-forward	0.04 s	1.68%

Fig. 12(d) compares the input currents generated by each controller. Note that all four controllers start supplying the maximum current for achieving fast heating and stabilize to zero, except for the PID controller which exhibits an oscillatory behaviour. In terms of power, it can be concluded that the feed-forward controller still offers the best performance.

Fig. 13 shows the total time response (heating and cooling) of the SMA under the feed-forward control. Note that the set temperature A_f is achieved fast and precisely while the reset temperature M_f is achieved exclusively by air convection. A 0.69 Hz bandwidth is to be expected for continuous cyclic operation.

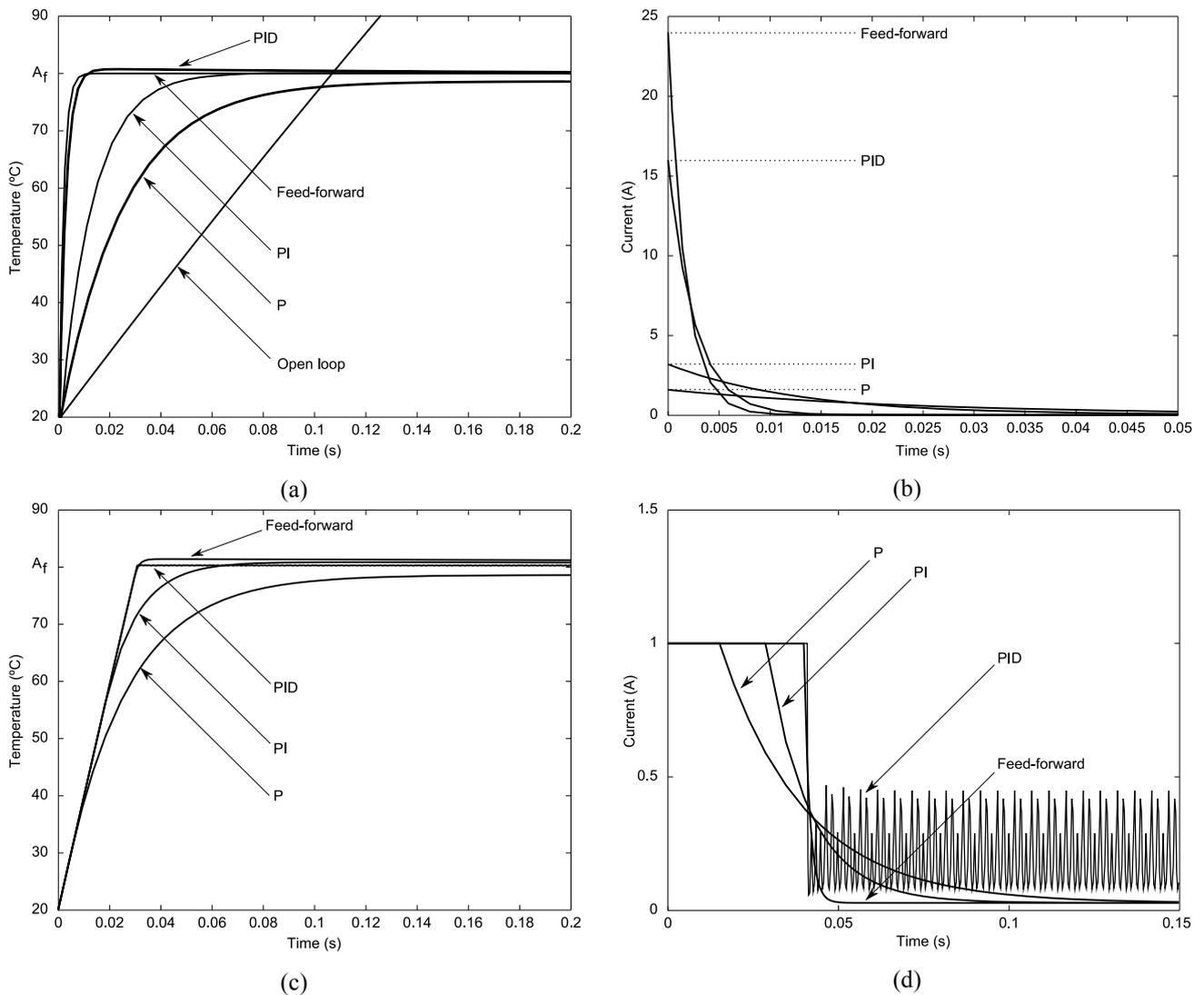


Figure 12 Dynamic responses in closed loop: (a) Improvement in step response using P, PI, PID and Feed-forward compensations and (b) the driving current inputs supplied to the SMA. (c) Comparison of step responses using a current limiter block. (d) Limited current inputs supplied to the SMA.

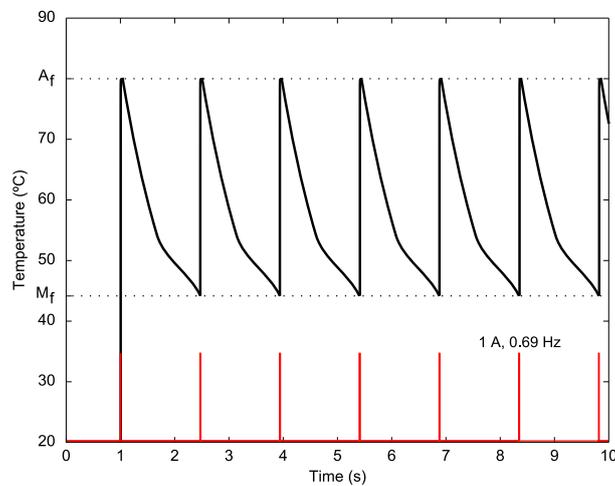


Figure 13 Total time response of the NiTi wire under the feed-forward control and free-air convection conditions.

5 CONCLUSION

This paper addressed the thermal behaviour of SMAs, especially the NiTi one, and presented a comparative theoretical study of the performance of a set of temperature controllers for improving their actuation speed.

A first step developed the thermal dynamics of SMAs through a simple analytical one-dimensional model and incorporated the non-linear phase transformation phenomenon which revealed a delay effect in the time response of SMAs. This delay effect is significant for cooling, however it can be reasonably neglected for rapid heating purposes.

A second step compared the performance of a set of linear controllers for improving the heating process: P, PI, PID and Feed-forward compensations were evaluated. To prevent overheating and thermal fatigue, these controllers take into account the maximum heating current at which NiTi wires can safely be heated. Feed-forward adds a new control approach that can significantly improve the thermal dynamics of SMAs.

REFERENCES

- [1] Liang C. and Roger C., One-dimensional thermomechanical constitutive relations for shape memory materials, *Journal of Intelligent Materials, Systems and Structures*, 1(2), pp 207-234, 1990.
- [2] Ikuta K., Micro/Miniature shape memory alloy actuator, *Proc. of IEEE International Conference on Robotics and Automation*, pp 2156-2161, 1990.
- [3] Raboud D., Faulkner M. and Lipsett A., Superelastic response of NiTi shape memory alloy wires for orthodontic applications, *Smart Materials and Structures*, 9, pp 684-692, 2000.
- [4] Haga Y., Mizushima M., Matsunaga T. and Esashi M., Medical and welfare applications of shape memory alloy microcoil actuators, *Smart Materials and Structures*, 14, pp 266-272, 2005.
- [5] Farias V., Solis L., Melendez L., Garcia C. and Velazquez R., A four-fingered robot hand with shape memory alloys, *Proc. of IEEE AFRICON*, pp 1-6, 2009.
- [6] Velazquez R., Pissaloux E., Hafez M. and Szewczyk J., Tactile rendering with shape memory alloy pin-matrix, *IEEE Transactions on Instrumentation and Measurement*, 57(5), pp 1051-1057, 2008.
- [7] Bellouard Y., Conception de dispositifs en alliage à mémoire de forme en microtechnique, *PhD Thesis*, EPFL, 2000.
- [8] Fu Y., Luo J., Hu M., Du H., Flewitt A. and Milne W., Micromirror structure actuated by TiNi shape memory thin films, *Journal of Micromechanics and Microengineering*, 15, pp 1872-1877, 2005.
- [9] Velazquez R., Hafez M., Pissaloux E. and Szewczyk J., A computational-experimental thermomechanical study of shape memory alloy microcoils and its application to the design of actuators, *Journal of Computational and Theoretical Nanoscience*, 3(4), pp 538-550, 2006.
- [10] Sutou Y., Omori T., Yamauchi K., Ono N., Kainuma R. and Ishida K., Effect of grain size and texture on pseudoelasticity in Cu-Al-Mn based shape memory wire, *Acta Materialia*, 53, pp 4121-4133, 2005.
- [11] Kafka V., Shape memory: a new concept of explanation and of mathematical modeling. Part 1: Micromechanical explanation of the causality in the SM process, *Journal of Intelligent Materials, Systems and Structures*, 5(6), pp 809-814, 1994.
- [12] LExcellent C., Goo B., Sun Q. and Bernardini J., Characterization, thermomechanical behavior and micromechanical-based constitutive model of shape memory Cu-Zn-Al single crystals, *Acta Metall. Mater.*, 44(9), pp 3773-3780, 1996.
- [13] Ortin J. and Planes A., Thermodynamics and hysteresis behavior of thermoelastic martensitic transformations, *Journal de Physique IV*, 1(C.4), pp 13-23, 1991.
- [14] Boyd J. and Lagoudas D., A thermodynamic constitutive model for the shape memory materials. Part II. The SMA composite materials, *International Journal of Plasticity*, 12(7), pp 843-873, 1996.
- [15] Tanaka K., A thermomechanical sketch of shape memory effect: one dimensional tensile behaviour, *Res Mechanica*, 18, pp 251-263, 1986.
- [16] Ivshin Y. and Pence J., A constitutive model for hysteretic phase transition behaviour, *International Journal of Engineering Science*, 32, pp 681-704, 1994.
- [17] Dubbel H., Beitz W. and Kuttner K., *DUBBEL handbook of mechanical engineering*, Springer-Verlag Berlin, 1995.
- [18] Dynalloy, Flexinol data sheet. Updated information available at: www.dynalloy.com
- [19] Potapov P. and da Silva E., Time response of shape memory actuators, *Journal of Intelligent Materials, Systems and Structures*, 11, pp 125-134, 2000.
- [20] Brailovski V., Trochu F. and Daigneault G., Temporal characteristics of shape memory linear actuators and their application to circuit breakers, *Materials and Design*, 17(3), pp 151-158, 1996.
- [21] Featherstone R. and The Y.H., Improving the speed of shape memory alloy actuators by faster electrical heating. In: M.H. Ang and O. Khatib (Eds.), *Experimental Robotics IX*, Springer Tracts in Advanced Robotics, pp 67-76, 2006.

VEHICLES HEADLAMPS GLARE EFFECT REMOVAL

Giorgio Ligios

Andrea Manuello Bertetto

Dipartimento di Ingegneria Meccanica, Chimica e dei Materiali
Università degli Studi di Cagliari

ABSTRACT

The presented work is the result of the development, realization and verification of a new device, designed to reduce and remove the glare effects of the automotive vehicle headlamps over the other drivers. Through the successful realization of a prototype it is possible to directly verify the fundamental principles of the device, by measuring the quality and the strength of its effectiveness.

Keywords: car, polarization, glare, safety, lighting

1 INTRODUCTION

The trend of the past few years in the automotive field is clearly bending towards the use of more and more powerful external lighting, high beam headlamps as well as dimmed headlights. This is also proven by the broadening use of plasma headlamps, sometimes labeled as HID (High Intensity Discharge), which emit gas at high pressure into the atmosphere. The use of the new technologies, which generate optical groups profoundly different from the halogen ones, generates new problems, such as glazing the drivers of the incoming vehicles. The proposed solution takes into consideration different perspectives, in order to lead to the development of new headlamps with high lighting power and no glaze effects over other drivers and pedestrians, who in particular will be made more visible in condition of poor external illumination. Focus is spent in improving the modern plasma headlamps safety, by integrating in them a new device with the function of radically reducing and even removing the reciprocal glaze effects over and from other vehicles.

The other very important function of that device is to enhance the external visibility in hostile conditions such as rainstorms, or presence of reflecting floods on the road coating. The headlamps technology has not improved for many years. One of the few milestones was achieved in the Fifties by the introduction of asymmetrical dimmed headlights to better illuminate the roadside. The roadside is the most critical part of the lane as it comprehends the shoulder of the street, whereas the other side is usually much brighter. A further milestone was achieved in the Sixties, when the new halogen filament bulbs substituted the old iodine ones. For the same amount of current consumption, the new bulbs guaranteed the double lighting power of the old bulbs, despite the one filament. Nowadays the halogen headlamps still constitute the most popular lighting system for automotive vehicles and civil applications. The structure of a common halogen headlamp is very simple. A metallic filament made of tungsten is disposed inside an ampoule made of quartz and containing noble gasses as iodine, krypton or xenon. The known Joule effect is the fundamental principle why and how a halogen bulb works. Electricity is led along the metallic filament, which warms up to about 2700°C. It becomes incandescent, therefore starts irradiating lighting radiation.

Contact author: G. Ligios¹, A. Manuello²

¹nexus.l@tiscali.it.

²manuello@dimeca.unica.it.

This technique is unfortunately very inefficient, as the 85% of the energy delivered to the bulb is transformed into heat and the remaining 15% only becomes visible light. Another restraint of the filament bulbs is its lifetime, as it commonly lasts between the 350 and 700 hours [1]. The “eyes” of the vehicles experienced no substantial modifications for decades, despite their limitations. They are fundamental safety devices though. They allow the steering despite the presence of fog and during the night, grant the other vehicles recognition, and can also be used to “communicate” with the other drivers, e.g. to warn about the presence of some danger, if correctly used. Taking into account that the 60% of fatal accidents occurs during the night, despite the reduction of the traffic down to the 35% [2], it comes obviously out that the vehicle lighting systems play a fundamental role in the people safety. The lighting system is the only way how to achieve a correct visibility of the external environment, when the road lightning is not enough. During the Nineties the revolutionary German automobile manufacturing companies BMW and Mercedes started offering the new expensive optional of xenon headlamps or plasma headlamps. This kind of bulb belongs to the group of gas emitting lamps. An ampoule made of quartz contains gas, in particular xenon and a small amount of mercury, and two electrodes made of tungsten (instead of the filament). Between the electrodes an electric arc is generated. It is only a few millimeters long, in order to more precisely focus the light. The normal tension given by the battery of the car of 12V is not enough to originate the electric arc, but some thousands of Volts (usually 15-20 thousands), so a trigger is used instead. When compared to a traditional headlamp, the light generated by a xenon bulb is usually very bright, it is similar to the sun light, and the light beam is broader and more penetrating (100 meters, instead of 60 meters of a halogen bulb).

The lifetime of a xenon lamp is about 3000 hours, therefore they usually do not need any kind of maintenance during a vehicle lifetime, when the halogen lights of a vehicle work for an average of 400 hours. At first, xenon headlights have been applied just to single high-beam, and then spread to the dipped headlight: this is the so-called bi-xenon projector. The transition between the two types of illumination is through a canopy that is being controlled by an electromagnetic device and allows you to vary the height of the beam by moving the bulb up and down inside the lamp, as if you had two separate filaments, one for high-beams and one for dipped headlights, like H4 lamps have. Both xenon lamps and halogen lamps can be equipped with a range of devices that helps guiding the light beam. Let's talk about dynamic and automatic leveling to prevent blinding the drivers from the opposite direction and the steering system, which is called cornering (see Figure 1.1), to better illuminate the inside of curves.

For xenon headlamps, dynamic and automatic leveling that replaces the manual one by the regulator on the instrument panel) is specifically required by regulations for approval of many countries: a series of sensors that are installed on the front and rear axles “read” changes in buoyancy of the vehicle due to acceleration, braking and static load, and they continuously adjust it via electro-mechanic actuators, the inclination of the beam, keeping it in optimal condition. The steering system can be realized through an additional ellipsoidal lamp that is mounted at a certain angle, which is fixed in the simplest cases or, for more complex ones, it can be dynamically adjustable (usually between 15° and 35°), relative to the direction of travel, and it can be also controlled software. The dynamic orientation, which is present in more complex xenon lamps, makes you able to adjust the light distribution to different driving conditions (i.e. city, country, highway, traffic), for left and right driving. For example, a wide and symmetrical beam is required for city roads, while having a deeper beam is preferable for driving in the highways.

The Italian law and EU regulations require that cars driving on our roads should be equipped with two front headlights, which must have each two lights in front projection to the direction of Forward, along with the yellow/orange indicator of the direction. These lights are called dipped headlights and high-beams. These last ones have a higher power than the previous. Their use during the day is governed by the Highway Code [3], but here we briefly mention that the dipped headlights are being used for normal driving in low light conditions, while those denominated high-beams are used in cases of outdoor very-low lighting. And here problems begin.

Because of their great power, high-beams cannot be used in all driving conditions because their “light jet” is strong enough to annoy those who drive in the opposite direction. So using high-beams while driving in rural roads and not when you cross other driver is a good idea, as well as the Driving Code says. In fact, every time you meet those events, you should always switch your dipped headlights on, as the name indicates as well, because those lamps have a adjusted intensity that does not blind those you meet while driving. This has been said according to the norm.

But we know that this does not happen, or at least it does not always happen and there are common situations where you look away from the street because you get dazzled by the high-beams, even quickly, with the consequent increased risk of getting in a car accident.

The situation gets worse if we consider the adverse weather conditions such as strong rain, as well as an imperfect cleaning of the windshield can be dangerous.

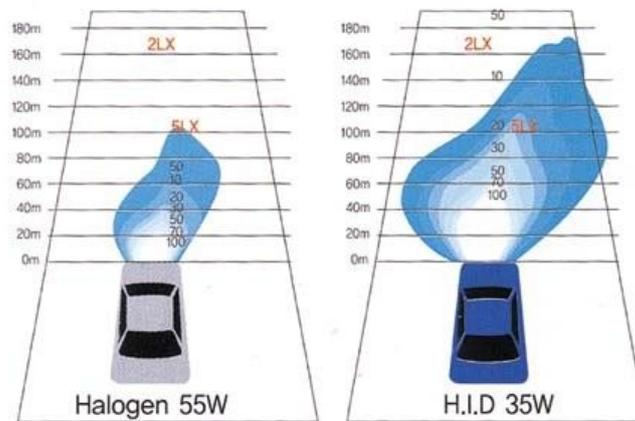


Figure 1 Different shape and distribution of the light beam of a halogen headlamp (left) and a plasma headlamp (right).
(Courtesy by Philips).

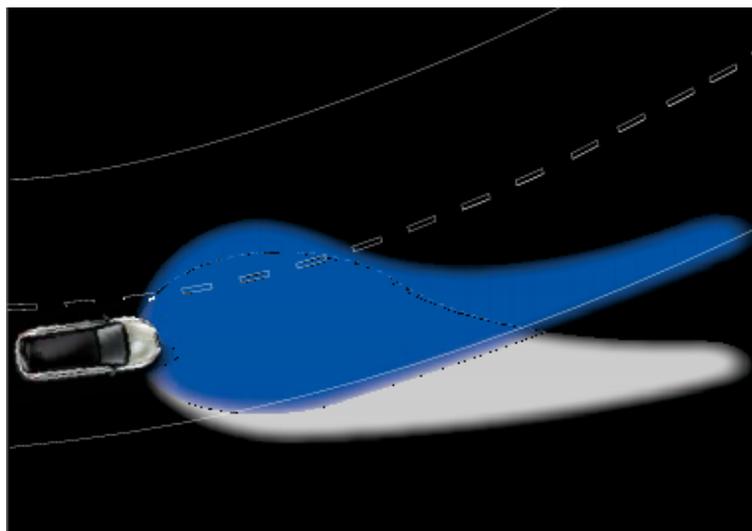


Figure 2 Example of using a beam steering cornering light to the inside of the curve.
(Courtesy by Fiat).

In these cases, dirt or water drops on the car's headlights will be spread, creating distracting reflections that would make the vision of the external environment very difficult, even if you're using windshield wipers. Before going on we must make a brief digression on the biological process of "vision", since this also involves subjective aspects such as increased or decreased personal sensitivity and other issues related to the type and nature of the used light's source. When a person moves from a normally lighted environment to an almost dark one, it gets blind initially and then gradually it regains vision. The adaptation is fast at first, but its completion requires at least 40 minutes to get done. This adjustment assumes an increase in sensitivity and it is due to three causes:

1) A chemical phenomenon, which is the regeneration of visual purple, also called as rhodopsin. This substance is contained in rods, it is altered by light and starts the stimulation of nerve cells. It regenerates quickly in a dark environment and its accumulation increases cells' sensitivity.

2) A nervous phenomenon, for which the individual ganglion cells (which are specialized cells within the retina) receive pulses from a greater number of rods, whose effects then increase. Because of a second mechanism, there is a decrease in visual acuity (which goes along with adaptation to the dark, especially in the fovea) and in perception of colors, especially yellow-orange-red.

3) An optical phenomenon, which is the enlargement of the pupil made by specific muscle fibers. As it is known there is a circular hole at the center of the iris, which is the pupil, through which passes all the light that is then distributed over the entire retina. The diameter of this hole is not constant: it widens up to 8 mm in diameter in the darkness and reduces down to 1 mm in bright light conditions.

In this sense, the pupil of the eye is comparable to the diaphragm of the cameras, especially those that are equipped with automatic exposure: the variation of its diameter is unconsciously controlled because of a nervous "reflection", on the basis of media environment's brightness. Moving from darkness to bright light, reverse phenomenon occurs: there is an initial period of glare, followed by a loss of sensitivity that after a while allows a correct view again. Adaptation to the light is faster than adapting your eyes to the dark, in fact it takes minutes for it to be complete. In addition to chemical and neural mechanisms mentioned above, here too there is an optical system that is more closely, and it is the automatic shrinking of the pupil.

Concerning the natural reaction to sudden changes of light, we see that the perception of a sudden increase or decrease of the amount of light in the eye causes an equally sudden change in the pupil opening, which narrows itself to protect the retina (this is a process called *myosis*), or it gets larger to increase the amount of light collected (this is a process called *mydriasis*). This biological mechanism is very fast and it takes only a few tenths of a second to happen.

2 THE PHENOMENON OF GLARE

When a bright light strikes the eye without allowing the pupil to adapt to the light radiation increasing (like it happens when you watch a flash or the headlights directly coming in the opposite direction) we get a temporary vision's trouble that is called "*glare*", which means that, besides the normal constriction of the pupils, we have a momentary weakening of the vision that leads to not see the contours of dimly lit objects very well.

This phenomenon occurs because of the radiation that reaches the retina completely consumes *rhodopsin*, which is the pigment included in retina's rods and used in the process of vision. Since there is a time interval of several minutes before the pigment is re-synthesized, the dazzled eye cannot see clearly for a while. Meanwhile there is the presence of "posthumous images" in the visual field, which are the same ones that can occur in seeing a brightly lit object (like a window pane in a bright day) and then turning the eyes on a dark background. your eyes, you will still see the first object, even if it will be very confused and pale.

This is because the chemical reactions caused by light in the retina do not vanish instantly during the light-dark transition. The posthumous images will then be "positive", which means that, if the stimulation of your retina will be short, those images will show the same light-dark objects in your vision. But if retina stimulation is prolonged, your retina may get tired and the posthumous image will appear with a reversed contrast. In case a luminance source significantly greater than the average of the actual or apparent source in the area of vision will fit in your vision field, the condition defined as *glare* will always cause a situation of disorder that can develop into a reduction in visual performance and come to a temporary loss of vision. This phenomenon leads to a reduction of the perception's speed and visual acuity, along with a abnormal increase in contrast ratio. Depending on the level of the phenomenon, it can occur a glare that causes discomfort and it is called *discomfort glare*, which is a psychological disorder, or else it can occur some glare effect that causes a real reduction of visual acuity (which is called as *disability glare*). Obviously, the discomfort's graduation is very difficult and it depends essentially by the activity that a person is doing, So the fact that an obstacle is being both visible or not under certain circumstances depends essentially by the sensitivity of the observer. This sensitivity is a function of luminance distribution in the visual field.

Therefore, if we examine the eye's behavior in detail, we can observe that, under normal circumstances, the sensitivity of a certain portion of the retina is being determined by the illumination of the same portion, and it is also being influenced by the illuminance on the other retina's parts. This phenomenon is commonly known with the name of *induction*. If distribution of luminance changes, the eye must adjust to the new light that has been created, and it takes a variable time to happen, which is called *adaptation time* [4]. Since we are not very willing to talk about the biological mechanisms associated with this phenomenon, I will just say that this occurrence is still a major cause for uneasiness in the person who is affected by it, and it causes an objective decrease in the distance perception as well as in view of obstacles by the driver. Another problem is related to the wavelength of used light, because the human eye is not equally sensitive to all frequencies present in the visible spectrum, which goes between 380 nm (violet) and 750 nm (red), but human eye greater sensitivity is limited to the green-yellow region and it also depends on the received intensity of light radiation. The *photopic* regime occurs in conditions of high-light intensity: cones in the center of retina mainly perceive the light, and the relative sensitivity gets its maximum at 555nm. The *scotopic* regime occurs under conditions of low intensity: the light is perceived primarily by rods at the edge of the retina, and in this case the relative sensitivity gets its maximum at 507nm.

Maximum sensitivity of human eye has been set at a wavelength of 555nm by convention, but if we deviate from this value we will see that sensitivity to electromagnetic radiation gradually decreases to zero at the extremes of the visible spectrum. This phenomenon shows that lights that have different spectrum (but with an equal intensity) are being perceived as a different power, and then ultimately they light in a different way. With all this said, we need to add the fact that the greater sensitivity towards green / yellow colors affects the adaptation of the pupil's diameter. So, for example, when a blue or violet light directly hits the eye of an observer, he perceives it with an apparently greater intensity but he sees yellow colors.

This feeling is because, while with green color the iris contracts in case of strong light, purple or another color that are at the edges of visible spectrum do not sufficiently stimulate the contraction of the pupil. The direct consequence of it is that the retina gets a greater amount of light, which creates a glare effect, although it is not the relative color for the maximum of its sensitivity, giving the impression of higher intensity than the counterpart of the greenish-yellow that has the same power. As final result, objects hit by light radiation get a lower illumination, and consequently the amount of collected details decreases giving a general visual discomfort due to glare effect. We can see that, in the case of night lighting by car's headlights we should take into account its specific wavelength, along with their power and the amount of provided light. All these considerations lead to think that the problem of outdoor lighting from cars should take account into headlights' pure light intensity and into the color temperature, or if you want, in those of main dominant. A blue-oriented light beam that has a blue dominant color in it, as we have seen it will be less visible and therefore less useful than a lower intensity light beam that has yellow as dominant color.

The current trend is, however, where many automakers are moving to, mounting xenon headlights on new cars, whose light has the dominant color in those spectrum's regions that have a higher frequency. These types of projectors are very efficient in terms of energy and light intensity, but they have got the disadvantage first explained to own a so powerful light beam that disturbs other drivers. More and more frequently it happens to get dazzled, especially in a city where, because of the ruined road, cars are subject to more bumps than on other roads.

These car "pitching" movements make the fact that light beam does not always remain parallel to the surface of the road, but it hits the drivers' face that drive in the opposite direction, causing them considerable discomfort. In addition, we can also see another problem related to this phenomenon while driving in a city, which is the reduced ability to perceive an obstacle (such as pedestrians) when you cross a car with other too powerful projectors on.

In these cases we turn away from lights and our pupil narrows, so a pedestrian who is in the middle two cars that are running in opposite directions between each other, becomes completely invisible even if it is simultaneously illuminated by two cars. This condition is not uncommon and indeed often causes serious accidents involving pedestrians, especially in rainy conditions when we use dipped headlights while driving in a town that has even a proper outdoor lighting. Here is where we give the best proposal for this issue, which is about an optical device that consists of two parts each with special filters, and whose purpose is to eliminate all the glare effects caused by the use of existing headlights. This would at least greatly reduce the discomfort caused by a direct light coming from cars running in the opposite direction, but it would anyway leave the ability of these lights to light the external environment intact. The principle on which it is based is the one of the light beam's linear polarization of the headlight, which goes through the use of linear polarizing filters placed appropriately within the car's headlight.

This device would therefore be an integral part of it and it would drive out only the component of light that has the polarization plane we wanted, shielding then the unwanted part. At this point the light beam would give up the projector in a completely identical way as to normal cars' lights, going then to illuminate the space in front of the car and you would not see any difference compared to a car equipped with standard headlights. However we would see a difference if we looked inside a car that comes in the opposite direction and has the second part of the device inside. This element would be constituted by a second linear polarizing filter in front of the driver's face, similar to the sun visor, which would have joints in it in order for it to be folded when its use is not necessary (see Figure 3).

A filter would be placed within it then, whose polarization axis is properly oriented to show an angle of 90 degrees, differently than the one installed inside the first car's headlights. This regulation of the two polarizers and the fact that the orientation of the polarizer inside the car would be like the one placed on their own lights, would not allow the driver to see directly the light emitted by the headlights of cars running in the opposite direction, but would normally make the driver see the external environment and its objects. This is because the headlights light coming directly to the driver's eyes should go through two orthogonally oriented linear polarizers.

And as we know from optics laws, there would be a lack of lighting radiation transmission in addition to the screens themselves. Regarding illuminated objects instead, there would be no problem, because the polarized light that is being reflected from objects' opaque surfaces would lose the characteristic of having a unique polarization plan, and it could still largely go through the polarizer placed in front of the driver's face.



Figure 3 Internal polarizer screen installed on a foldable support in the sun visor set.chart.

The end result of all this is that the driver would continue to see the outside through the light produced by his own lights and also by others, while he might not be dazzled by other motorists' headlights. The amount of their emitting that reaches the eyes of drivers in the opposite direction would not be a problem, but if it gets greatly reduced to an intensity like the one emitted by tail lights, it would show a color close to blue and purple.

This color is justified by the fact that polarizing filters that are being used for tests have a cine-camera origin, specifically they are a Manfrotto® brand with Polaroid filters, used for astronomical photography. They have a good quality, although they got a function field that is more efficient in the green spectral region but they fail to polarize as well the frequency of shorter wavelength. There are still good polarizing filters on the market that are working good in all regions of the visible spectrum. It must be also said that, even if some light made by headlights is still visible, this would not cause any disturb but instead it would be useful as a parking lamp. Tests have in fact shown that this condition is very useful since the low residual purple brightness helps to better identify the car's contours and do not disturb drivers' vision. So lack of glare effect would prevent the pupils to excessively contract, allowing the eye to capture more details of the objects illuminated by a dim light.

A real Polaroid® leaves a $K\%$ fraction of the electric field's component to pass parallel to the axis of transmission and absorbs almost all of the electric field's component (which is perpendicular to the axis of transmission) by transmitting less than 1% of polarized light in perpendicular direction. The Polaroid-type polarizers are identified by the "NH X" initials, where X is the percentage of incident intensity that is being transmitted, where a non-polarized light is present. An ideal polarizer would be called as "NH 50", and it is the one that passes completely to the electric field's component that

goes parallel to the transmission axis and absorbs the perpendicular axis.

The most common Polaroid® systems are the ones called "NH 32". For these $K = 80\%$ Polaroid® the intensity of transmitted light is 32% of the incident one. This is because a certain percentage of the incident light (which is up to 4%) is reflected from each of the two sides of the lamina. The transmitted light (almost completely polarized in the direction of the optical axis) can reach a maximum of 40% un-polarized incident light intensity. The Polaroid's polarizers are excellent over the entire visible spectrum, even if they have a 100% absorption in the region of blue color.

3 INCREASING BRIGHTNESS

As we have seen, problems start when we consider the conversion efficiency of light by the polarizing filters.

As we know from the optics laws, this is limited by the physics of the process, which sets a theoretical maximum percentage of luminous intensity output by a polarizer to any value by 50% [5] of the incoming one.

This happens because the normal light emitted by a source, whatever a lamp of a projector is, does not have a defined polarization's plan, and then, once this gets fixed by the passage of light flow through a polarizer, the component of radiation with parallel electrical axis' plan allowed by the latter rises. This would mean a real loss of light intensity produced by the headlights, which at best would be a 50% loss. According to current Legislation, regarding the maximum luminous flux that can be emitted from cars headlights[6], we can see that maximum allowed values are showing up to 1650 lumens for "high-beams", and 1000 lumens for those "dipped headlights". Just for the record, we have that all light bulbs approved in Europe by a voltage test of 13.2V have the following nominal values for emission of light:

H1: 1550 lumen +/- 15%

H3: 1450 lumen +/- 15%

H7: 1500 lumen +/- 10%

Considering those tolerance, these lamps cannot emit higher lumens than the following values:

H1: $1550 \times 1,15 = 1782$ lumen max

H3: $1450 \times 1,15 = 1667$ lumen max

H7: $1500 \times 1,10 = 1650$ lumen max

Using headlamps that emit the highest legally permitted values, with an efficiency of the polarizers that has been arbitrarily chosen and is equal to 80% (those on the market also come to overcome that percentage), the use of filters would then bring the total luminous flux output respectively on the values of:

$1000 \times 0,5 \times 0,8 = 400$ lumens for dipped headlights

$1650 \times 0,5 \times 0,8 = 660$ lumens for high-beams

Such values are obviously not enough for getting a right outdoor lighting, and the device that is simply added to the headlights would be useless. In fact, the increased visibility of obstacles against the light and the lack of blinding by drivers would not compensate for such a high loss of light from the headlights. There are the two ways to overcome this problem: a) increasing the luminous flux produced by the headlights, in order them to offset the losses; b) using optical devices and additional measures to reduce the amount of light lost through the polarizers.

The first way is certainly the most simple and straightforward. Standard halogen lamps, which are marked as "H4" (60/55 watts), are currently being installed into clusters. The indication of power gives the values for power consumption that are related to the two bulb's filaments, both the high beams and dipped headlights. The symbol "H4" indicates the type of lamps used and one of the two most common standard models installed. Another widely used model is the "H7" type, where, instead of a single bulb that inside has both filaments, there are two independent bulbs, each with its own filament, and each one is installed in a dedicated reflector inside the headlight (there are also other standards for lamps, such as H1 and H3, but we will not mention them here, unless in the tables of power).

With this last solution, the parable is for high-beams as well as dipped headlights, while the different light beam shape (which is higher and deeper for high-beams, lower and angled for headlights), is defined by a particular design of the parable itself. This parable performs its different focus function because the lamp's filaments appear to be set in two different focus points.

So, passing from one light to another, this allows us to get two jets of light shaped as we want. As for the light beam emitted, below there is a table with data models and best known brands on the market. (See Table 1) The percentages shown above refer to the tolerances for the real lumen than the product actually rating. Those values are provided by the Manufacturers.

As you can see there are already several models that are installed on regular cars, even though they show higher values than the maximum allowed. These values would still be not enough for our needs, which is to recover the loss we get due to the use of polarizers. You should then turn your attention to categories of lamps that are called "special" because they have a special range of use, even if they are available in the market.

This is the case of boosted lamps, which have higher absorption than the standard ones and are being used for applications such as motor sport rally or as additional lights for off-road vehicles. It must be said that these lamps are usually found in sales, and we can find common models with different characteristics and light output, but their use is still illegal for normal road applications. Unfortunately, fans of "tuning" (which means an extreme customization for motor vehicles accessories) use those kind of lamps we mentioned above very often. The light power emitted by these lamps often gets to be twice the allowable values, and in this case, glare effect would be a very serious problem.

There are indicated values for some lamps currently on the market just to give an idea what luminous flux values they can reach to. As we can see from the lists above, standard H4 type can be used in the majority of vehicles in accordance with the existing rule, and it is only the fifteenth place in the list on the emitted luminous flux.

Such table shows only a part of that type that is already on the market, because every day there are more and more powerful lamps. The last two entries of the table deserve a separate discussion. They are the two Osram® D2S and D2R lamps, these belong to that type called HID (High Intensity Discharge), which are made by an electric arc into an xenon atmosphere. They are also known as plasma lamps. Their function is quite different from the normal filament bulbs one, and it is based on an arc-discharge that is maintained between two tungsten electrodes in a bulb, which contains xenon gas under a high pressure (3.5-10 Mpa) and mixtures of elements from the "rare lands" group which are in variable proportions [7].

The light they produced is greater than any other type of lamp, and it has got several advantages: •The light efficiency is far bigger than the halogen light one, it reaches about 80lm/w against to 20lm/w of standard H4 lamps type, or the 31.3 lm/w for the 80W Ocean® strengthened Halogen lamps.

They have a higher "color temperature" that reaches 4500 Kelvin, raising up to more than 6000 K against the 3200 K for halogen lamps, having a tone so much closer to the sunlight's one. Luminous flux values are ranging from double to triple than the most powerful halogen lamps values, since new "HID" lamps start from 3200lm and go up 6800lm in dual bulb configuration.

Table 1 - Lamps H4

Philips Vision Plus	1895/1150	lumen	max
Philips Premium	1815/1100	lumen	+4/-15%
Osram Silvestar	1700/1075	lumen	+7/-7%
Osram Super	1700/1075	lumen	+7/-7%
Osram Allseason	1700/1000	lumen	+10/-7%
H4 standard (USA)	1665/1000	lumen	+15/-15%
Osram Cool Blue	1650/1000	lumen	+15/-15%
Philips BlueVision	1650/1000	lumen	+15/-15%
Osram Light@day	1650/1000	lumen	+15/-15%
Philips WeatherVision	1650/1000	lumen	+15/-15%
Philips Longerlife	1650/1000	lumen	+15/-15%
General E EuroBlue	1650/1000	lumen	+15/-15%
Osram normale	1650/1000	lumen	+15/-15%
Phoenix	1650/1000	lumen	[for.all.H4]
H4standard EUROPEAN	1650/1000	lumen	+15/-15%

Table 2 - Lamps H1

Phoenix	130W 3400	lumen	+/- n.d.
Ocean	130W 3120	lumen	+/- 15%
Osram	100W 2600	lumen	+/- 10%
Philips Rally	100W 2500	lumen	+/- 15%
General Electric	100W 2750	lumen	+/- n.d.
Ocean	100W 2000	lumen	+/- 15%
Phoenix	100W 1850	lumen	+/- n.d.
Ocean	80W 1750	lumen	+/- n.d.

Table 3 - Lamps H3

Phoenix	130W 3500	lumen	+/- n.d.
Ocean	130W 3200	lumen	+/- 15%
Ocean	100W 2900	lumen	+/- 15%
Phoenix	100W 2600	lumen	+/- n.d.
Ocean	80W 2500	lumen	+/- 15%
General Electric	100W 2350	lumen	+/- n.d.
Osram	100W 2300	lumen	+/- 10%
H3 standard ECE-R37	1450	lumen	+/- 15%

Table 4 - Lamps H4 (dual light)

Philips Rally	100/90W	2900/1700	lumen	+/- 15%
Osram	100/80W	2900/1700	lumen	+/- 15%
Phoenix	130/90W	2880/2280	lumen	+/- n.d.
Phoenix	100/90W	2400/1900	lumen	+/- n.d.
Phoenix	100/80W	2400/1900	lumen	+/- n.d.
Phoenix	100/55W	2400/1000	lumen	+/- n.d.
Philips Vision Plus	n.d.	1895/1150	lumen	+/- n.d.
Philips Premium	n.d.	1815/1100	lumen	+4/-15%
HID Osram D2S	35W	3200	lumen	+/- 15%
HID Osram D2R	35W	2800	lumen	+/- 16%

Table 5 - Approved Lamps (hours)

Narva LongLife	1250
Philips Long life	500/1250
Osram Longerlife	600/1200
Osram normale	400/650
Ocean normale	400
Narva normale	400
Philips AllWeather	150/400
Philips BlueVision	100/350
Philips Visionplus	200/350
Narva Xenon B.L.P.	50/350
Narva Xenon B.L.P.	30 350
Osram Allseason	150/350
Osram Silverstar	150/350
Osram Super	150/350
Osram CoolBlue	150/250
Philips Premium	100/250
General Electric	225 [14V]
Phoenix 55W	225

Table 6 - Powered Lamps (hours)

Ocean 80W	200
Ocean 100W	150
Ocean 130W	150
Phoenix 100W	100
General Electric 100W	100 (14V)
Osram 100W	50
Phoenix 130W	25
Philips 100W Rally	20/50

Required electric power for 35W to those that reach 3200lm brightness, while we have seen that a normal H4 lamps series has 55/60W power consumption respectively for dipped headlights and high-beams, with a value of luminous flux not higher than 1650lm.

This means less heat emission in the optical unit inside, and less wear issues for all the reflective surfaces within it. One obvious solution to deal with less brightness due to the use of circular polarizing filters would be to replace normal bulbs with an equivalent type of lamps, the plasma ones. For example, if we consider using a baseline model like Osram® D2S ® that has 3200 lumens (with a polarizing efficiency of 80% and 50% cut of brightness due to the very functioning of the polarizer) we would see that the final values would be:

$$3200 \times 0.5 \times 0.8 = 1280 \text{ lumens}$$

This is a lower value than the optimum one indicated by the standards, which is 1650 lm.

However, we need to consider that the given values come from tests that have been performed on new lamps. As operating hours increase, emitted brightness by halogen tends to sharply decrease. After 150 hours of use on an average life span that rarely exceeds 400 hours, values get lower down to 85% than the initial ones, getting around 1400 lm brightness. As an example we report life span H1 bulb-type, which are listed by duration that has been measured in operation hours. (*See Table 5 and 6*).

This reduction in brightness during the useful life does not happen with "HID" type of lamps, where the light output remains almost constant throughout the lamp's life span, which duration now commonly exceeds the 3000 hours for. The loss of just over one hundred actual lumen, which are made by HID 3200 lm lamps and polarizing filter, does not become excessive if we consider light with normal use that have lower flow values than nominal ones, also because basic model plasma lamps with not too extreme emission values have been used as a reference.

It would be enough if we consider plasma arc lamps with values around 3500 lumens to get the same flow conditions as shown below:

$$3500 \times 0.5 \times 0.8 = 1400 \text{ lumens}$$

This is a perfectly reasonable value for using polarizing filters, because the loss of just a hundred lumens would be offset by a reduced effect that comes from a visual impact of lights on other drivers. Remember that such device must increase the driving safety as its ultimate goal, by eliminating glare effect to other drivers and increasing visibility in artificial lighting conditions.

What has emerged as a result after making "road tests" made us see that detected faint glow in the remaining lights with filters made cars more visible to other drivers. Similar to what happens when we use only the tail lamps, the optical unit no longer appears as a single bright "light globe", but the arising low light purple that eliminates the "backlight" effect allows us to easily distinguish cars' contours, making it much more visible. We have been also able to verify that this happens in cases of reduced outdoor lighting and this proves it even more useful in judging distances by other means. This possibility is crucial when trying to value the speed and drivers' positions while they are driving in the opposite direction.

4 POLARIZATION CONVERTER

We got so far just considering the increasing the light output of the headlights to compensate the 50% lost by passing through the polarizers as a solution. That loss would then be offset by using more powerful lamps. However, we can also put a device capable of recovering a fraction of the radiation light lost in passing through the filters along to lamps' replacement, making it by converting part of the light that could not be polarized along the axis we wanted, and that has got the wanted polarization's plan, which would then go to be added to the useful one.

This process is achieved with the use of so-called "*half-waved sheets*", which are special optical devices capable of rotating the incident radiation's polarization plan of 90° angle and which are going to transform, for example, horizontal axis' linearly polarized light into linearly polarized light by vertical axis. We can know their function if we consider a dual-refracting crystal plate with flat and parallel faces material, such as calcite, which a non-polarized light beam affects to. The incident wave is being usually separated into two components, which are the "*ordinary*" and "*extraordinary*" radius and are linearly polarized with perpendicular vibration plans due to its crystalline structure, which has different refractive indices in different directions of light rays' propagation, making

then a speed difference for propagation of the two components.

The axis that defines the direction with higher refractive index is conventionally called "*slow axis*". Instead, that axis which defines the direction with lower refractive index is called "*fast axis*". This is the one where the optical axis of the crystal lies on. This is because if you choose the first polarized wave so that the electric field oscillates in a direction that is perpendicular to the optical axis, that polarized wave crosses the plate with a speed as $V(\perp) < C$, and it will re-emerge from the back having the same polarization. The second polarized incident wave will get its electric field oblique to the optical axis, and one of its perpendicular components will go with a speed as $V(\perp)$; another component that is parallel to the optical axis will go with a speed as $V(\parallel) \neq V(\perp)$.

Then the result we get is a deviation of the wave's direction (called refraction), which will rise from the back having an identical polarization to the original one it will be set in a different location than the first considered wave.

Ultimately, a dual-refracting crystal separates a natural light beam into two light beams that are linearly polarized and have the vibration plans going perpendicular to each other. We can build optical devices with special properties using dual-refracting crystals, such as the so-called "*late plates*", which are really important in the area of polarization applications. A late plate is composed by a dual-refracting material that has optical axis' parallel surfaces.

If you carved a perpendicular linearly polarized monochromatic wave, you will produce a phase difference between the component wave with electric field that is parallel to the optical axis and the one with the electric field that are in perpendicular direction, and this happens because of the different speeds of propagation. The two light beams will come out of the plate as overlapped, with a phase shift that depends on the thickness of the plate, and they will recombine together, producing again a polarized wave, but in different way that the incident one.

In the *half-wave plates*, radiation's wavelength and plate's thickness make the displacement as equal as an angle that is like half of the wavelength of optical path difference between the two outgoing components.

The component that goes in the direction of propagation (which is relative to the "slow axis") rises a middle-period later than the other component, and their recombination still gives a linearly polarized wave, but now it has got a rotated vibration plan. In the specific case of a 45° angle between the fast axis and the direction of vibration of the electric field of polarized incident light beam, we will get an outgoing polarized light beam that is orthogonally placed on the original one.



Figure 4: Summary table of the various forms of standard lamps and comparison of some features, gathered by type of bulb. (Courtesy by Lamp.com)

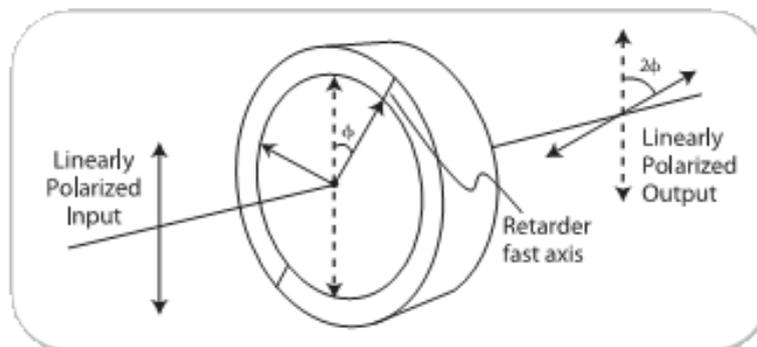


Figure 5: Half-wave Plate. (Courtesy by Meadowlarks)

The "half-wave plates" are very common devices and their use is now part of commercial devices that work with polarized light. This is the case of the projectors that are being used in professional and now increasingly also as a hobby. In these devices, the image generation is based on using a polarized light beam. Once it is focused and made pass through an LCD colored matrix, it is projected on an

outdoor screen that is specially crafted. Recent technology developments and the need for more bright headlights, led to creating a polarization device that is separate in two distinct parts, due to the coupling of crystal polarizing beam-division, with a separation of 90 between the two beams (see Figure 5)

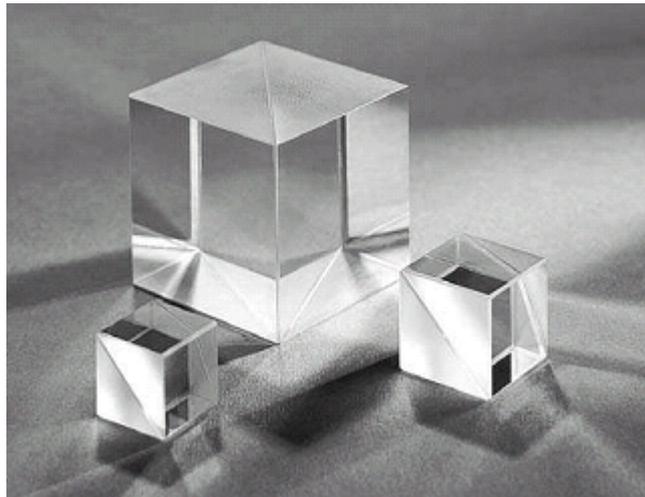
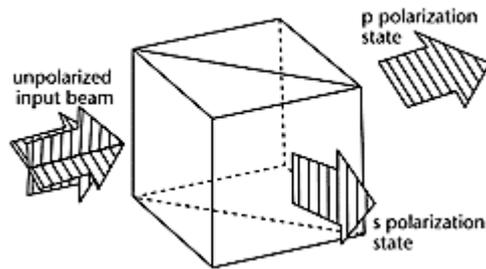


Figure 6.

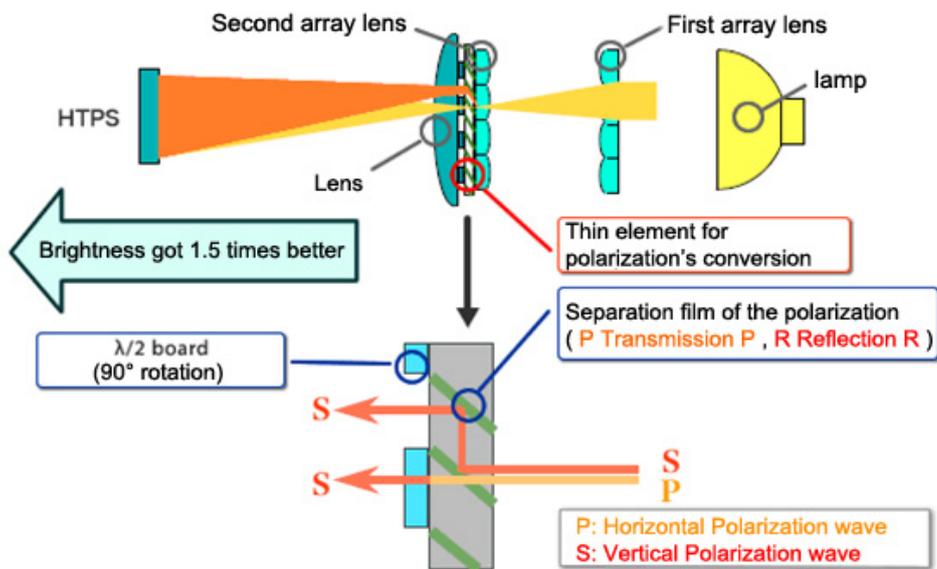


Figure 7: Polarization's conversion element. (Courtesy by Epson)

A *half-wave plate* has been placed in front of each crystal, which is located on the crystal's front surface. This is to turn the front status of the outgoing radiation's polarization from "P" to "S", and make it equal to "S", which is the beam coming out from the side. This Epson® made-by device has been miniaturized and integrated into its two components, achieving a single panel for it which an array of polarizer crystals and half-wave plates have been made on. All elements have been placed in parallel rows and were joined together to form a square area, through which the light of projector's lamp passes (*see Figure 7*).

As you can see, the lamp's light goes first through a square array of lenses and then it goes through a second lens array to get to the polarizing device itself. Here the two linearly polarized components' separation occurs, according to the "S" vertical axis and the "P" horizontal one. The first one, rising from the crystal with a 90° angle on to the light beam axis, reflects itself on a film of separation that allows only the "S" component reflection. After that it gets out of the crystal as a polarized light, according to the vertical axis. Instead, the other "P" component is not being turned away from the crystal and it keeps going along the axis until it gets to the "half-wave plate" that is placed on the exit face. Then the polarization's rotation that transforms its state from "P" to "S" occurs within it. You can see how, with such a device, a portion of the light that would be lost is recovered and could be used by the projector. This leads to a significant increase in brightness, as stated by Epson® itself, and it is estimated as a 50% increase compared to the luminous flux output without using this device. This technology could then be easily used for the realization of the proposed device, which will have polarizing filters installed in, just like projectors described above have. Once such improvements have been adapted to different needs (such as the elimination of the first series of lens and a change in the size and number of matrix cells) using them would lead to a loss of light smaller than the model that only has polarizers into. The luminous flux output would no longer be penalized by 50% , but it would stop at much higher values, reaching theoretical values that would reach a 75% of brightness produced by the lamp. In fact, as explained before, you should consider the efficiency of the polarizers, but the fact that such a device would be able to increase the total device's brightness to a 1.5 point remains. This would perfectly balance the brightness loss, even combined with the use of enhanced lighting, and it could reach even higher values of light output without causing any annoyance to other drivers.

5 THE PROTOTYPE

The second solution mentioned above is certainly extreme if we take account of the still high cost and the necessity of using proprietary patented devices. With this said, we chose to use a more conservative approach in order for us to run tests on this issue, following a philosophy about costs reducing and device's complexity.

We came to adopt such an increase in headlights' light output through the use of plasma arc lamps (HID), along with using standard polarizing filters that are being used also in cine-camera's works, which are easily obtainable on the market and their prices are definitely lower than devices made by Epson®. This allowed us to build a working prototype on which light intensity measurements were made, light that is emitted in standard conditions along the use of circular polarizing filters. The used device is made by a pair of car headlights (made by Fiat®), which have been installed in second-set "Fiat Punto®" cars.

This model is now obsolete, but the type of headlights is still being commonly used on many other cars and it was chosen because of easier market's retrieval. These projectors are as "*lenticular*" type, which means they have parables for dipped headlights that are installed behind a front glass lens whose task is to better restrict and direct the light beam along the roadway. This is a solution that has been widely adopted by all car manufacturers, since it reduces the light beam dispersion and enables you to focus it on desired areas with ease and efficiency that cannot be obtained by using traditional parables. The used headlights have H7 lamps type installed in, which have a single filament into. Dipped headlights and high-beams are separated and have two different side-by-side installed reflectors, where each one has its own H7 lamp. This allowed us to work more easily on dipped headlights, without having to make corrections on the second parable, since size and external diameter of the one dedicated to high-beams were excessive for the retrieval of polarizing filters that could have an appropriate size.

The headlight has been modified by removing the transparent front section to get access to the dipped headlight's parable, whose front is closed by a glass converging lens (as you can see in photos). A support with a circle removable crown has been then installed on it, and a fixed linear polarizing filter's metal ring has been glued on its top. This has been made to allow the second polarizer ring to freely rotate, in order for it to get a right viewing angle for the polarization axis. The used filter consists in fact of two crowns. The first one is threaded and fixed, the second one is mobile and able to rotate at 360°. This feature is due to the polarizing filter's particular choice, which has a cine-camera origin, (made by Hoya®) and it is "PL" type, with an internal diameter of 82mm. The entire filter crown has been fixed to the headlight with fast-release metal hooks. This solution has allowed us to perform different tests , using and not using a filter on the same optical unit. A device has been then made next to the filter's crown, which is also removable, motorized and connected by a wheel gear to the polarizer's external ring. Its purpose is to make the filter rotate on itself and ensure that the headlight's light can be linearly polarized along an axis which you can freely choose the position for, having all intermediate positions between the vertical and horizontal directions available.

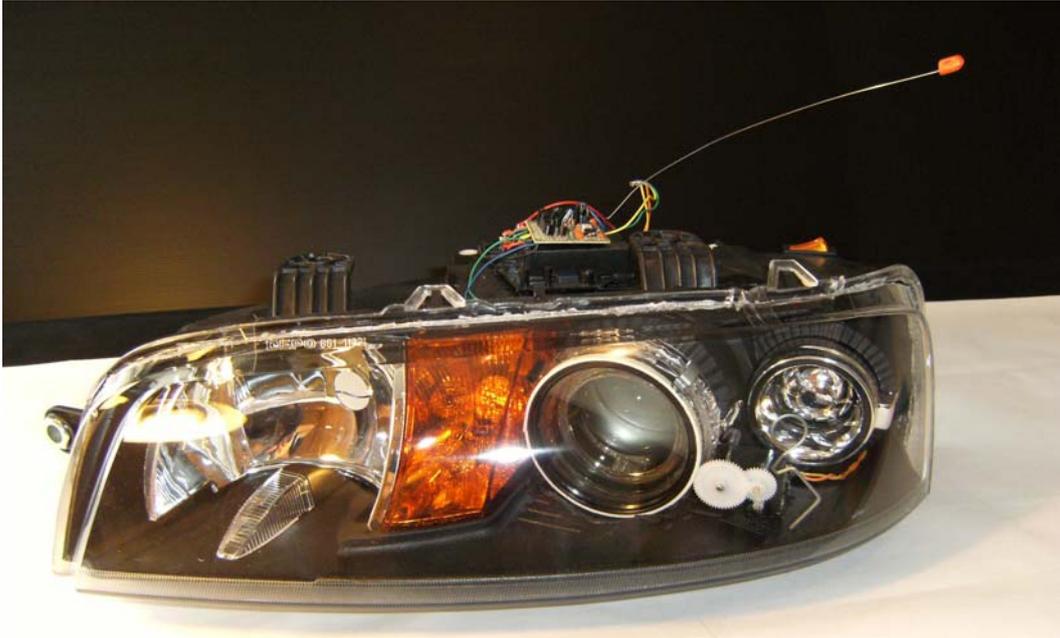


Figure 8: Used headlight.

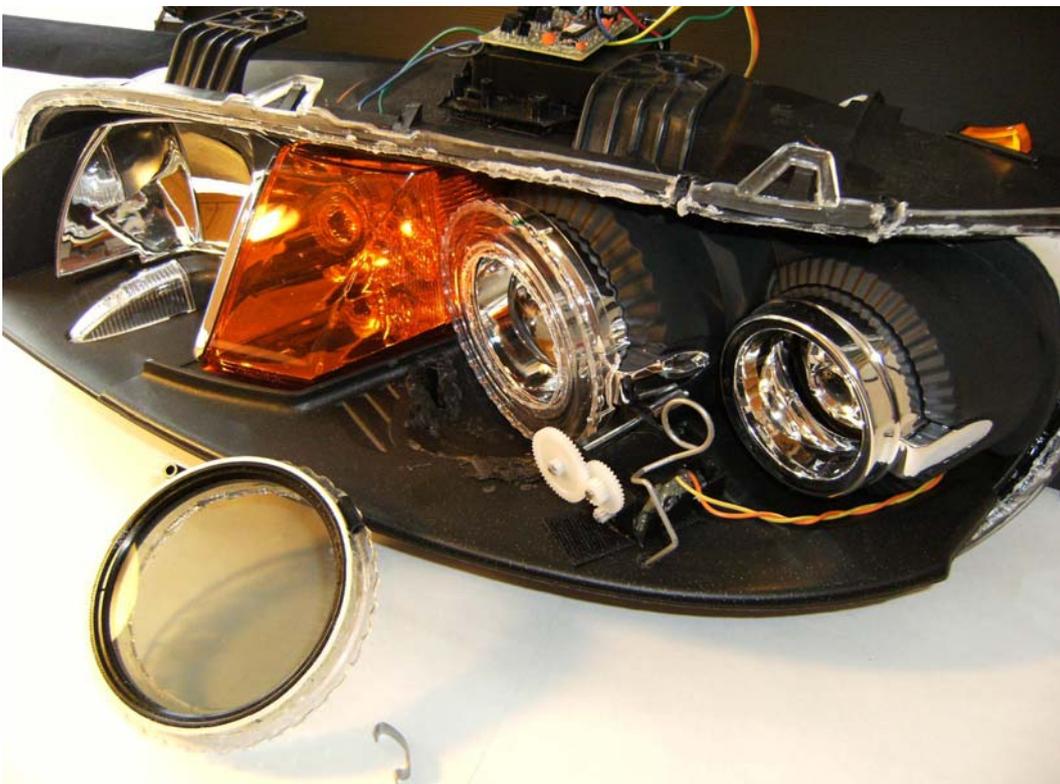


Figure 9: Headlight without the transparent cover and the polarizer, which has been removed from the dipped headlight's body.

The device's control and power have been taken off the car engine, and are being managed by an electronic circuit that is powered by batteries and works by a remote controller. The entire mechanism and the radio module have been assembled starting from the engine and gear ratio of a battery powered toy, and then they have been modified to make them suitable for our purpose. Operating voltage has been lowered from 4.5 to 3 volts, so you can rotate the metal ring slowly enough get more control of the position.

The reasons behind using these materials are many.

First, a cost reduction of the prototype, in order for us to demonstrate the device's low cost. Then there is the possibility to use a ready and tested control system just as the remote controller, without having to redesign it. A third reason is the possibility to perform various alignment of the axes of the "closed headlights" filters, without having to reopen it every time to realign the polarizers according to the desired angle. In addition to this, "targets" have been fixed on the polarizer's crown in order for it to have a visual indication of the polarization axis' position at any time. This solution was necessary since radio remote control was digital: on / off, without the possibility to precisely install it using position sensors, according to the desired angle. To power the headlight we used a 12 Volt / 7AH lead-acid accumulator, whose power was enough to shortly manage two 55W bulbs' power consumption, or to power just a single lamp for longer than enough to complete all tests.

We have initially thought to use a power supply resulting from a 220V power supply, but the needs of portability have led us to choose Pb battery. All electrical connections have been treated so as to have a minimum impact on the optical unit, getting out the most of existing connections and minimizing the changes in order to have the entire spotlight as similar as to an unmodified one.

Everything has always been inspired by a perspective of "pre-existing modification", which is the stated purpose of the project. Standard H7 55W, Philips® brand halogen bulbs have been used for the reference tests without filters, while xenon HID bulbs have been used for powered lamps. Supply section and assembly kits are made in Germany, but the bulb is still made by Philips®.

The two xenon lamps' kit we have used to develop our prototype contains two shielded high voltage power supplies that are needed for lamps function. Since, they are lamps that discharge gas in the atmosphere, the HID bulbs need an operating voltage of 85V and require a trigger to the ignition that has a voltage higher than 20,000 V. Data on plates declare that they reach 3400 lumens light output and a color temperature of 4200 Kelvin, with a life span of 3000 hours.

Below are the specifications issued by manufacturers:

As can be easily seen from the specific data (*see Figure 11 and 12*), xenon lamps would be rule-out [6] because of their power and color temperature, so they could not be installed in cars without violating the Highway Code [3].

In fact, its luminous flux values are much higher than those generated from the halogen lamps that have not polarizing filters (which greatly decrease the amount of emitted light), so they would be very harmful to other drivers.

If you directed those lights to a driver view who drives in the opposite direction, you would risk to "blind" him and to cause a high risk level of getting a car accident.

6 INNER SHIELD

The second part of the device, which is the one that needs to be installed inside the car by the sun visor, was simulated by using a pair of plastic glasses with a Plexiglas screen. On its hole's edges was added a guide so that polarizing filters could be fixed on it, according to the desired orientation. We have deliberately chosen not to produce the prototype engine, since it was unnecessary to understand the functioning of the device and the data collection from tests. Filters alignment, according to the desired axis, has been made manually by removing and reinserting the two polarizing filters directly into the glasses edges.

Regarding the two polarizers shape (one each eye), we chose the square one because it is easier to put it on its stand, and because if we're working in a controlled environment we can make observations while keeping the glasses always in horizontal position. Going forward, once we chose an orientation degree for polarizers, whether vertical or horizontal, we could keep that axis without getting any "rolling" movement as it happens in cars, which would continuously change the angle of polarization plane. The alignment engine function for polarizers would be to orient themselves according to the preset angles and to maintain this orientation even if the car's order changed. The idea behind the device is actually the extinction of light from a car's headlights that runs in our opposite direction, made by the combined effect of polarizing filters placed on its headlights and those that are located on the sun visor within your car.

We see that if filters do not have a perfectly orthogonal angle between the polarization axis, they cannot decrease the headlights amount of light, so a portion of light will however get to the driver' eyes, causing him more visual discomfort due to its glare effect.

So, the filters powering (as it has been made on the prototype without the aid of any tilt sensor) would ultimately continue to equalize the polarizers position, cancelling car's rolling motion in the case of a curve or in case of some lateral tilt on the road due to its particular conditions.



Figure 10: Xenon lamp with its power supply, before being inserted in the parable

Business Center Automotive Aachen	Product Information Sheet H7 12V 55W Blue Vision	No.: 725-1 13-Aug-1999																										
<p>Halogen Bulb H7 12V 55W Blue Vision Cap: ¹⁾ PX26d Type: 12972 BV ECE R37 E1 22Z Code No. 9240 210 171.. Made in Germany, Aachen <small>¹⁾ according to EEC 61-1 sheet T084 5-1</small></p>																												
<table border="1"> <tr> <td>d</td> <td>1,24 nom.</td> </tr> <tr> <td>e</td> <td>25 ± 0,2</td> </tr> <tr> <td>f</td> <td>4 ± 0,2</td> </tr> <tr> <td>h1(0°)</td> <td>0 ± 0,15</td> </tr> <tr> <td>h1(90°)</td> <td>0 ± 0,15</td> </tr> <tr> <td>h2(0°)</td> <td>0 ± 0,25</td> </tr> <tr> <td>h2(90°)</td> <td>0 ± 0,25</td> </tr> </table>		d	1,24 nom.	e	25 ± 0,2	f	4 ± 0,2	h1(0°)	0 ± 0,15	h1(90°)	0 ± 0,15	h2(0°)	0 ± 0,25	h2(90°)	0 ± 0,25	<table border="1"> <tr> <td>a1</td> <td>a2</td> <td>b1</td> <td>b2</td> <td>c1</td> <td>c2</td> </tr> <tr> <td>d + 0,3</td> <td>d + 0,5</td> <td>0,2</td> <td>0,2</td> <td>4,6</td> <td>4</td> </tr> </table>	a1	a2	b1	b2	c1	c2	d + 0,3	d + 0,5	0,2	0,2	4,6	4
d	1,24 nom.																											
e	25 ± 0,2																											
f	4 ± 0,2																											
h1(0°)	0 ± 0,15																											
h1(90°)	0 ± 0,15																											
h2(0°)	0 ± 0,25																											
h2(90°)	0 ± 0,25																											
a1	a2	b1	b2	c1	c2																							
d + 0,3	d + 0,5	0,2	0,2	4,6	4																							
<table border="1"> <tr> <td>Technical data</td> <td>R37</td> </tr> <tr> <td>Test voltage</td> <td>13,2 V</td> </tr> <tr> <td>Power</td> <td>56 W max.</td> </tr> <tr> <td>Luminous flux</td> <td>1450 lm +13% -7%</td> </tr> <tr> <td>Lumen maintenance</td> <td>85 % min.</td> </tr> </table>		Technical data	R37	Test voltage	13,2 V	Power	56 W max.	Luminous flux	1450 lm +13% -7%	Lumen maintenance	85 % min.	<table border="1"> <tr> <td>Rated life at</td> <td>13,2 V</td> </tr> <tr> <td>- characteristic Tc</td> <td>350 h min.</td> </tr> <tr> <td>- B3</td> <td>200 h min.</td> </tr> </table>	Rated life at	13,2 V	- characteristic Tc	350 h min.	- B3	200 h min.										
Technical data	R37																											
Test voltage	13,2 V																											
Power	56 W max.																											
Luminous flux	1450 lm +13% -7%																											
Lumen maintenance	85 % min.																											
Rated life at	13,2 V																											
- characteristic Tc	350 h min.																											
- B3	200 h min.																											
<table border="1"> <tr> <td>Maximum operating temperature:</td> <td>800 °C max bulb temperature;</td> </tr> <tr> <td></td> <td>400 °C max pinch</td> </tr> <tr> <td>Operating position:</td> <td>horizontal</td> </tr> </table>		Maximum operating temperature:	800 °C max bulb temperature;		400 °C max pinch	Operating position:	horizontal																					
Maximum operating temperature:	800 °C max bulb temperature;																											
	400 °C max pinch																											
Operating position:	horizontal																											
Property of Business Center Aachen - Germany		Page 1 of 1																										

Figure 11: H7 lamps type Specifications (Courtesy by Philips)

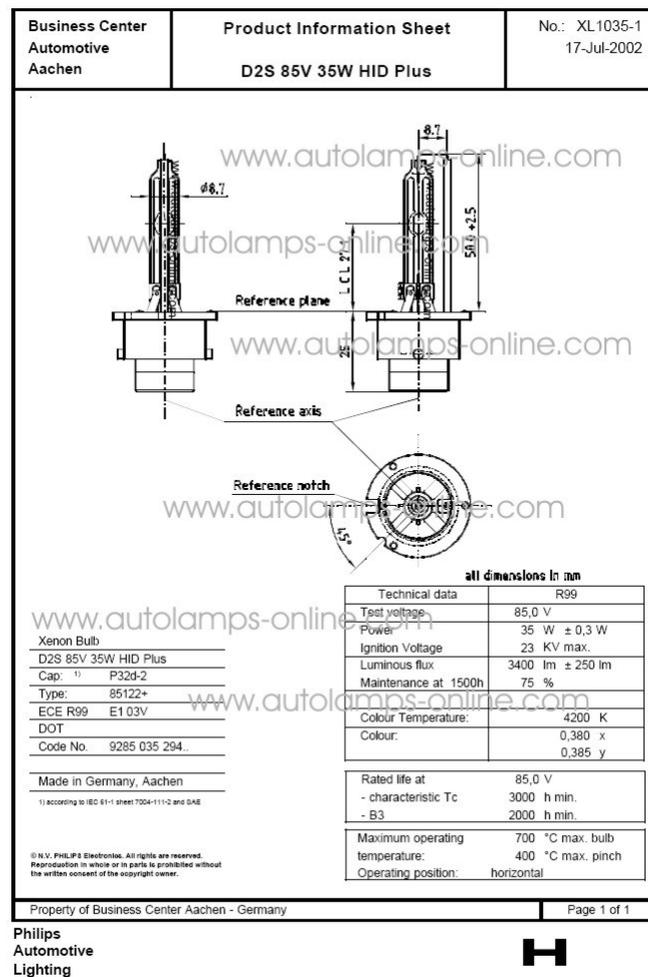


Figure 12: Xenon lamps Specifications (Courtesy by Philips)

6 MEASUREMENTS MADE

Once we have created a prototype, we went to see what differences we could find between the brightness provided by the normal halogen bulbs and the one made by plasma lamps. In order for us to obtain results that could be compared with each other and could have a general validity, we decided to set up tests that would provide us a vertical white screen setting and an optical that was perpendicularly placed to that screen, still aligned with its central point. The distance between headlight and surface was set equal to 3 meters, and 6 points have been identified on the screen, both taken in those regions with the highest light concentration as in those where headlamp's light beam has less brightness. This difference in brightness we have found is not an imperfect alignment of the reflector lamps and does not defect in curvature, but is entirely intentional. The reflecting surfaces design is actually so designed in order for it to direct the light beam preferentially to the right region and more upward than the

left side. If we think about driving a car we can understand what mentioned before, because it is much more useful to see the obstacles on the right side of the road than object coming from its left side. In fact, from this side we have the other side of the road which is an open space, and if we limited the light in that direction we would not cause any harm to drivers that are running on the same street. All this is achieved by limiting the light beam's height that goes out from dipped headlights through the parable design and with a metallic screen placed immediately behind the filament that is inside the lamp. In other cases like ours, where we used a lens headlight, this "cut" to the top of the light beam is obtained with a vertical aperture inside the parable, just before the glass lens. Together with the same ability to concentrate the light, this makes the output light beam to illuminate only from a street level up to the center of the dipped headlights lens' height.

We can prove this by looking at the light beam's shape that is being projected on the screen, where a very sharp horizontal line is separating the lighted area from the dark one. In the pictures below you will see a faint colored line at the edge between dark and bright area.

This is due to light refraction on the inner diaphragm. Once we had set a distance, we decided to directly measure the light levels on the screen, which has 6 fixed points in all the various test configurations of the headlight.

Besides these values, we have also made two further readings on a point placed two meters away from the screen and fixed to a height from the ground which made the selected point not being within the cone of light, which has the lens at the top of the headlight and the screen as a base.

This disposition left the headlight completely shadowed, which is a very useful condition since illumination's measure was not found out by measuring how much light got to the screen (with the sensor oriented towards the light projector), but how much light got to that point after being reflected on the screen itself. This result has given us the opportunity to value the amount light emitted from the optical group that goes to the observer's eyes after being reflected on an object, which actually is the white screen mentioned above. This is the more important result we got, because we think that an observer would be interested in knowing how much headlights' light would be able to come to his eyes and then, ultimately, how much he would be able to then see as an object. At this exact point, indicated by a "P" letter, we have made four readings for each of the eight tested configurations on clusters. Referring to lamp and polarizing filter, we will have the following lamps:

- Halogen and without filters
- Halogen and vertical axis filter
- Halogen and horizontal filter
- Halogen and 45° filter
- Xenon without filters
- Xenon and vertical filter
- Xenon and horizontal filter
- Xenon and 45° filter

The four readings we have seen in the "P" explain why we first measured the value of brightness at that point without putting a filter between the sensor and the screen, and then we measured again the brightness amount placing a linear polarizer between the sensor and the screen and performing three readings in horizontal, vertical and 45° direction.

The reason that has led us to perform only two readings instead of four is that the three registered values were completely identical when we placed the polarizer before the sensor and the screen. So we know by optics laws that a linearly polarized light loses its polarization state when it is reflected by an opaque non-metallic object.

In this way, whatever the state of light's polarization incident on the screen is, once this is reflected by the screen it completely loses this feature, showing characteristics that are equal to "natural" light. Then, once the light goes through the polarizer in front of the sensor, this would polarize that light and it would measure only the attenuation of typical luminous flux of the process, which, is however independent by the polarizer's orientation.

The reading on this point is therefore that one we would see inside a car, watching from behind the polarized screen that is placed under the sun shade.

In order for us to measure the light levels at various points we used a digital light meter, an LX-1330B model made by GainExpress® with fully adjustable scale, using a maximum resolution of 0.1 Lux, a silicon photodiode with maximum sensitivity set at 550nm and measurement frequency about 2/sec. The measuring ranges that we got are as follows:

Table 7

Measuring Range	Resolution	Accuracy
0-200 Lux	0.1 Lux	± (5%+2d)
0-2,000 Lux	1 Lux	± (5%+2d)
2,000-19,990 Lux	10 Lux	± (5%+2d)
20,000 -100,000 Lux	100 Lux	± (5%+2d)
100,000 -200,000 Lux	1000 Lux	± (5%+2d)

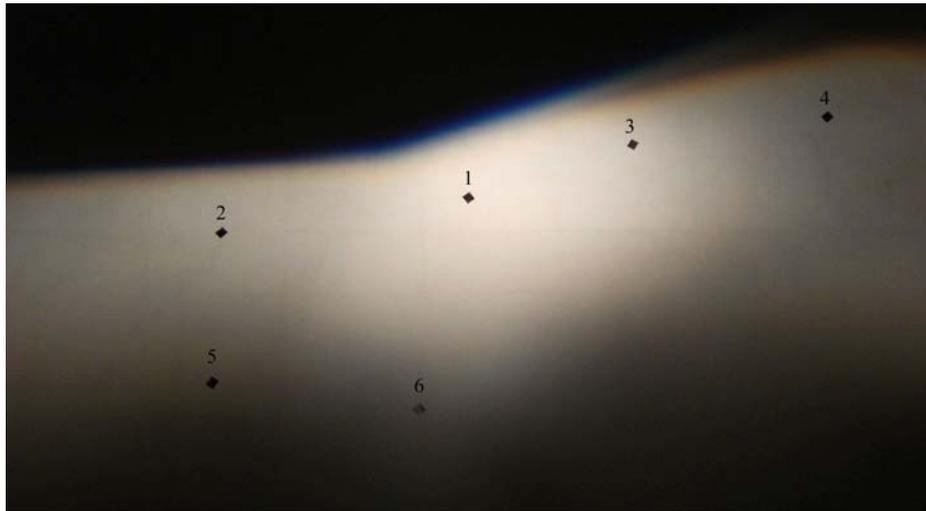


Figure 13: Reference points for Halogen lamps without filters

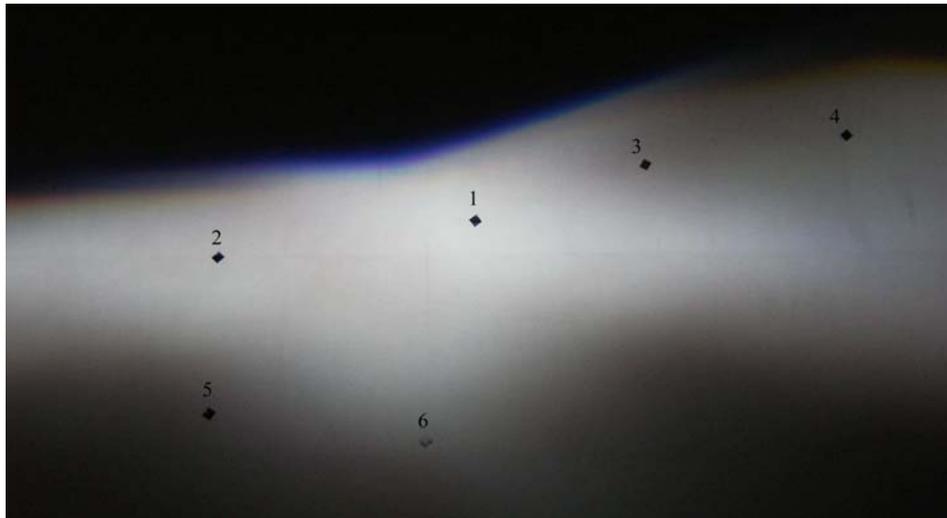


Figure 14: Reference points for Xenon lamps without filters.

7 RESULTS

As you can see from the photos of the screen (see Figure 13 and 14) there are points on it where measurements of illumination were made during the eight test configurations, and for each measurement we obtained three readings, considering the average of three records as the final result (see Table 8).

As we can see from the data above, we have obtained what we had proposed in the beginning, which was to obtain a system that would allow us to use polarizing filters without affecting brightness along with the ability to light the environment by the use of car's headlights. Going to look better at the results we clearly see that the illumination values provided by plasma arc lamps are more than twice the halogen lamps values. We see how values can have

being reduced for 50% by using circular polarizing filters, going to get to levels we had found in the halogen lamps configuration that were had no filter installed into. In fact, although values look very similar between each other, there is some discrepancy regarding No. 2 and No. 3 points.

They are about 200 Lux different than the initial configuration and that those lamps with xenon and filters installed into (see Drawing 1).

We cannot say that this happens because of a non proper functioning of the polarizers, but it is because of a different configuration of the HID's light beam.

As we earlier mentioned, tests have been made according to an alignment that would respect the original shape of an unmodified light beam, and although we have obtained a result, the shape of xenon bulb's support has created two

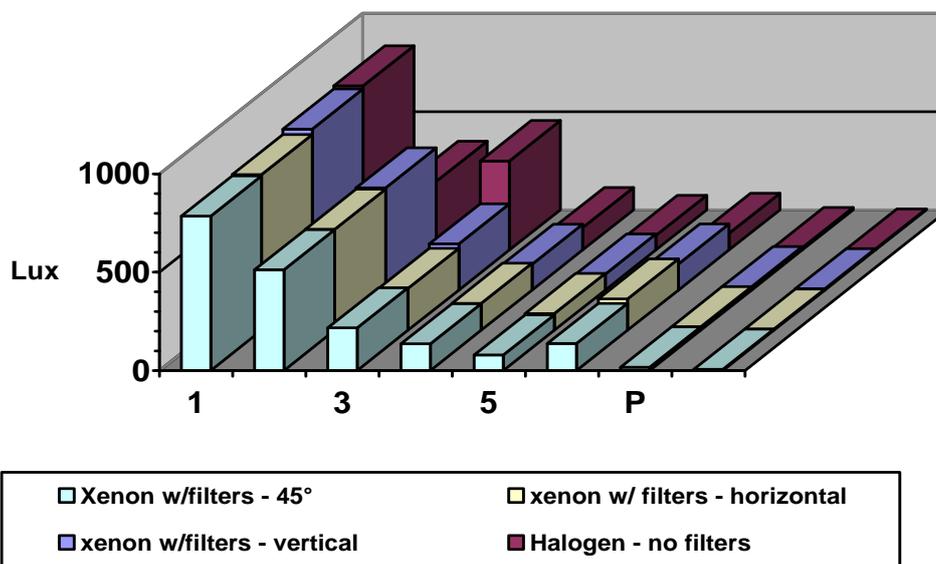
small areas into the field which this requirement has not been exceeded to.

As we can see by pictures, we see how the light footprint "design" in both cases (halogen lamp and HID lamp) is almost identical. We notice only a difference in point 2 and 3 (on the second configuration), where the area appears to be less lighted than the surrounding area.

This was due to a thin ceramic coating inside the lamp plasma's bulb, whose function is to isolate one of the two power poles from getting touched by the bulb itself, and then to create a shadow zone that is detectable when the light beam goes on a screen.

Table 8

point num.	Halo no Filter Lux	Halo.w/filt vert.Lux	Halo.w/filt ori.Lux	Halo.w/filt 45°.Lux	Xeno no filter Lux	Xeno w/filt vert Lux	Xeno w/filt ori Lux	Xeno w/filt 45° Lux
1	835	330	331	330	2002	820	792	786
2	350	130	122	124	1287	520	514	513
3	454	182	173	177	530	235	214	217
4	114	50	47	50	345	132	138	135
5	75	27	26	26	206	81	86	79
6	85	29	33	33	298	133	160	138
P	13	5.9	5.8	5.6	42.3	17.8	17.5	17.3
P.with Polar	5.5	2.2	2.2	2.3	16.2	7.2	6.2	6.2



Drawing 1: Comparison between the illuminations recorded in basis points with four standard configurations.

In general we can then say that what we got by those results is what we expected to get, even if we just look at the differences between xenon lamps' luminous flux values and halogen ones. Much less obvious for us was to confirm that if we were in a real situation with light beam projected onto a screen. If we consider only the lumen values expressed that came out from data we could never be sure that two lamps or two clusters can generate the light amount we were looking for on a same surface.

This is because one thing is the light emitted from the lamp and another is how it is concentrated by the parable on a more or less restricted area. For example, if a lamp which itself did not have great lighting values, was fitted with a parabola that concentrates all its light on a very small area, we would get very high Lux values if we measured that lamp's brightness in that area. On the other hand, if we had a very bright light but its parable "opened" its light beam without focusing it very well, we would get very low light

levels even with higher values of luminous flux than the previous case.

It is for this reason that tests were performed maintaining the same optical unit and ensuring that plasma arc lamps were generating their light output at the same point in the parable that was previously occupied by halogen lamps, which means that both lamps were located in the reflector's focus. In the following pictures (see *Figure 15 to 18*) we see how the device is really able to improve the visibility and breaks almost entirely down car lights glare effect that come to the observer's eyes.

Pictures were made in a real situation environment, which means that we were standing two meters far from it, in front of the headlight unit, directly framing it. The first image was obtained without using any type of filter, you can see how that picture literally "shoots" light to us, which means that we see only a fuzzy ball of light produced by the headlight's light. In the other picture we can instead notice what result we would get if we were using both linear polarizing filters, both in front of the headlight and in the front of the observer's face, with this last one orthogonally oriented than the first filter.

This one, as mentioned above, would be represented by the screen inside the car. From that picture we can see that brightness reduction from the headlight lets us see its

contours, even if light output is not completely blocked, and some brightness goes through the polarizers because of their not so good quality. By the way, in those pictures we see that light that still comes out of the headlight, despite the presence of filters, seems to be very large.

This is due to the short distance between the camera and headlight. This short distance is the worst situation that someone might find in reality, but we deliberately have chose that to show how it is possible to see the full optical group even in a condition like this, although the optical group is completely against the light.

The ability to see objects against the light (such as a pedestrian) was highlighted during tests using a "puppet" that was like a hypothetical pedestrian between two cars running in opposite direction, both with lights on. Under normal circumstances, we know from experience that the combined effect between headlights high brightness and absence of adequate external lighting would make this situation dangerous for a pedestrian. If our eye gets struck by the headlights light, it would suffer a violent glare effect, the pupil would then contract and the already outside dim light would not be enough to see an obstacle in front of us, with the result that you can create a very dangerous situation for your driving.



Figure 15



Figure 16

As you can see from the picture, headlights' brightness reduction from cars that we would meet would not generate any glare effect, so our eyes should not need to adapt again to new visibility conditions, and all this would make the

light coming from our car headlights enough for us to see an object on the road. We would have a very similar condition even if other cars wouldn't be present on the road and we had only our headlights on lighting the road.



Figure 17: Headlight with xenon no-filters lamp, with a hidden by backlighting object.

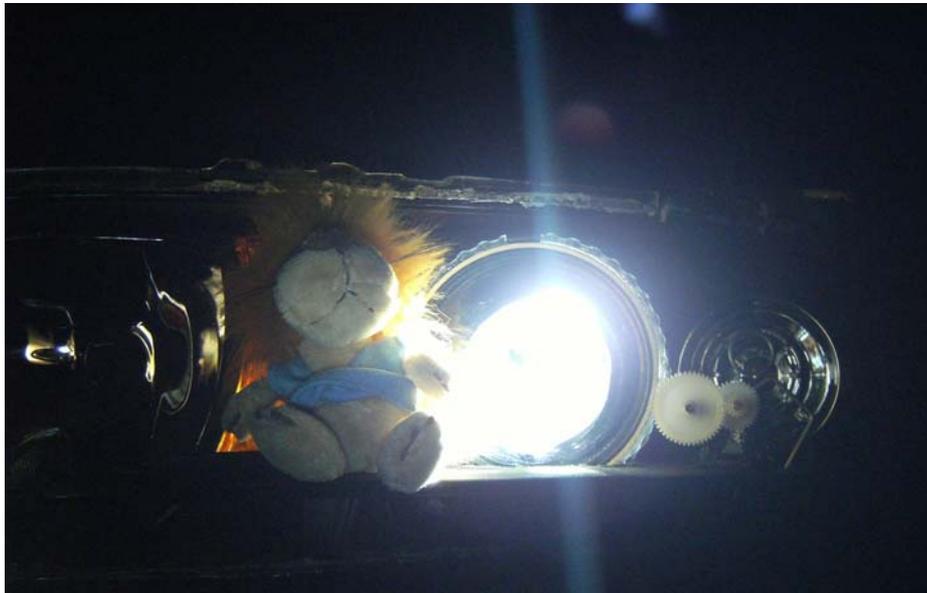


Figure 18: Headlight with xenon and filters installed, both in front the lamp and in front of the observer. The object is visible.

8 OTHERS APPLICATIONS

A further interesting filters application showed then their power during the prototype's development.

The device that has been designed to work when both drivers have the same device installed, so when they cross each other's road, filters cancel the other car's light. But

there is a case in which the device is also useful when used alone, which is when it eliminates unwanted reflections produced by our headlights light. All of us in fact may have been in a situation of such strong rain that we could not see even a few meters far from our car. In such event, when our headlights are reflecting their light on the water drops, it is

not rare that light creates a kind of mirror effect that makes us see the rain as a sort of "water wall", beyond which you cannot see anything. But if we oriented the optical unit polarizer according to a certain axis (let's say a vertical one) and those into the screen in a perpendicular way than the first, we would only see a depolarized light coming to us from a reflection on an opaque object, while all that light emitted by our headlights would not lose its polarization

state and would not be able to go through the inner screen. Here its usefulness is. The water drops, by their nature, act as tiny mirrors and are not able to depolarize the light coming upon them, not even when it affects their outer surface or when, after entering into the drop, it is being reflected from its rear surface on its air-water separation limit.

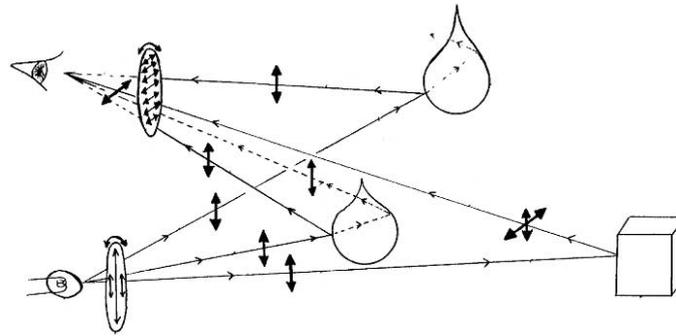


Figure 19

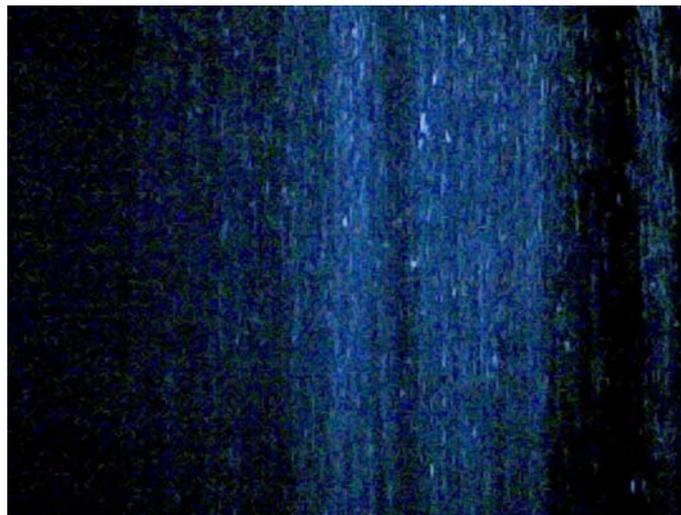


Figure 20: Rain simulation, seen without filters.

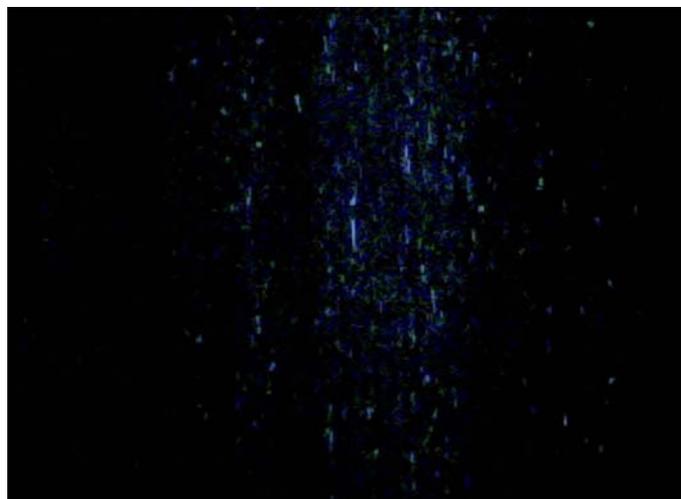


Figure 21: Rain simulation, seen using filters.

The final result is that the water drops seem to almost "disappear" from our view because headlights light that is being reflected from them cannot go through the polarizer that is inside the car. You can see this aspect on the two pictures above (see Figure 19 and 20), where we tried to make what we have previously explained, simulating the presence of rain drops in order for us to create a strong rainy situation. Unfortunately those pictures do not make that phenomenon so clear, because visual impact during strong rainy weather is much higher to the naked eye. Unfortunately we could not properly make the difference between headlights with filters and without them in case of rain, since the camera we used, although of good quality, does not behave like the human eye does, and exposure times were always too short to appreciate what we described.

9 FUTURE DEVELOPMENTS

From all we said before, it is clear that the proposed mechanism has more than a few advantages that would be of much help in improving road safety by increasing the visibility for drivers and for pedestrians walking on a roadway. The leveling device we have talked about earlier, which work would be to always keep the polarizers aligned at an angle to the ground, whether horizontal or vertical, could also have another function. Since it can be motorized, it could then be connected to sensors that are able to detect the amount of light that is going through the inner shield to the car. Doing so, it may make some adjustment like, for example, switching into a "rainy" mode, since it seems that tests had got a better "visual feedback" when the polarizer was positioned with an horizontal axis to a vertical position from that axis itself. That position is the best to prevent reflections by puddles from spreading around.

As for deciding about which angle we may choose if we wanted to use the device according to the original idea, it is not be possible to agree with the other cars or change the angle every time a car is coming from the opposite side.

Therefore, a solution may be found by a configuration that is good for all drivers and for both street's directions.

In order for us to do so, filters on optical groups must be counterclockwise turned than the vertical axis, at an angle of 45° , while internal optical groups keep staying oriented at 45° from the vertical axis, but in a clockwise direction.

With this configuration we can see that all cars headlights could keep this orientation, because, once they cross a street, they find themselves facing each other and the angle between the respective filters is always 90° .

The moving device for motorized filter would still have a reason to exist because it could always keep a constant angle of 45° from the ground. Between many positives things, there is also a downside that the prototype has shown during testing. As we know, most of the transparent plastic materials have the ability to depolarize the light.

We will not discuss here about dual-refraction on certain materials caused by mechanical and thermal-mechanical stress. We just need to remember that front headlights are made in plastic material. It goes without saying then, fixing the polarizers inside the headlight, including lamps and outer transparent cover, would be totally useless because the polarized light that comes from the filter would be immediately depolarized by passing through the coating.

In order for us to solve that problem, we should prepare a new cover design that would provide for a glass "window" inside the front transparent plastic part of the headlight.

This would maintain current safety levels provided from plastic material's shells, along with the possibility of using the illustrated device, giving clear advantages for the two systems integration. Once the device is integrated into the headlights, we could possibly think to also increase those limits imposed by headlights light emission rules. Those limits have been established in order to avoid glare effect to other drivers. But this would no longer occur if we could carry a solution like we said before into effect, allowing us to increase the emitted light output and get the benefit of improving visibility of obstacles and other cars.

Using circular polarizing filters would lead to the disappearance of the usual gesture to lower the headlights whenever a vehicle crosses another (gesture that unfortunately is not always respected), and there would be a great improvement in driving at the night, or at least in situations with poor outdoor lighting, such as heavy rainy weather condition or a situation of not perfectly clean windscreen. We should note that the device is still a simple and cheap mechanism, and if we built it in series, we could install it as an integral part of the car. Concepts and component parts would not even need special studies because, although they should be further adapted to the searched goals, they would be already existing devices used in conditions similar to the ones of the proposed equipment. The only problem would be to get deal with the manufacturers for using their patented devices, and possibly to implement a collaboration to come up with useful variations to make the device even better.

REFERENCES

- [1] Philips, *Philips Automotive Lighting*, www.automotive.philips.com, 2010.
- [2] ISTAT, *Rapporto Incidenti Stradali*, pp. 8, 2008.
- [3] *Nuovo Codice della Strada*, Art. 151, Art. 152, 2003.
- [4] Moncada Lo Giudice G., *Illuminotecnica*, pp.141-155, 2007.

- [5] Bennett J.M., *Polarization, Handbook of Optics*, M. Bass, McGraw-Hill, Vol. I, Cap. 5, 1995.
- [6] Norma ECE-R37, Rev. 3, 2008.
- [7] Moncada Lo Giudice G., *Illuminotecnica*, pp. 74, 2007.
- J.A. Dobrowolski, "Optical Properties of Films and Coatings", Vol. I, Cap. 42;
- D. Clarke, J.F. Grainger, *Polarized Light and Optical Measurement*, (Pergamon Press, New York, 1971);
- D.S. Kliger, J.W. Lewis, and C.E. Randall, *Polarized Light in Optics and Spectroscopy*, (Academic Press, San Diego, Calif., 1990);
- W.A. Shurcliff, *Polarized Light: Production and Use*, (Harvard University Press, Cambridge, Mass., 1966);
- D.A. Holmes, "Exact Theory of Retardation Plates", *J. Opt. Soc. Am.* 54, 1115 (1964);
- P.D. Hale and G.W. Day, "Stability of Birefringent Linear Retarders (Waveplates)", *Appl. Opt.* 27(24), 5146 (1988);

BIBLIOGRAPHY

University of Michigan:" Visual effect of blue-tinted tungsten-halogen headlamps bulb", <http://dmses.dot.gov/docimages/p62/133155.pdf>

J.M. Bennett, "Polarization", *Handbook of Optics*, M. Bass, (McGraw-Hill, New York, 1995), Vol. I, Cap. 5;

A DRY FRICTION MODEL AND ROBUST COMPUTATIONAL ALGORITHM FOR REVERSIBLE OR IRREVERSIBLE MOTION TRANSMISSIONS

Lorenzo Borello Matteo D. L. Dalla Vedova

Department of Mechanical and Aerospace Engineering - Politecnico di Torino
Corso Duca degli Abruzzi 24 – 10129 Torino, Italy

ABSTRACT

The motion transmission elements are generally affected by dry friction which may give rise to reversible or irreversible behaviour of the whole system. So, the potential high effect of the dry friction on the dynamic behaviour of the mechanical system requires proper simulation models characterised by high computational accuracy, nevertheless compactness and efficiency. Aims of the work are:

- the proposal of a general purpose physical and mathematical dry friction dynamic model and, consequently, the detailed description of two related numerical algorithms, developed in different computational environments (low level language and Matlab-Simulink), able to simulate the behaviour of a general dynamic system affected by dry friction and equipped or not with ends of travel;
- the simulation of some representative actuation runs in order to validate the proper accuracy of the mechanical device computational algorithm and the analysis of the results.

Keywords: coulomb, friction, actuator

1 INTRODUCTION

A typical mechanical problem regards the actuation systems in which one or more motor elements drive one or more mechanical users via a motion transmission; in some cases, the actuation system is a part of a servomechanism, expiring the control of position of the user. The simulation of the dynamic behaviour of these systems may require mathematical models having the ability to take into account the usually undesired effects of dry friction forces/torque, more or less affecting all the working conditions. If the system is equipped with mechanical ends of travel, their effect must be properly taken into account by the model

itself, without any prejudice of the correct dry friction simulation. Generally, whichever the motor or user types are, the motion transmission consists of a certain number of shafts, gears, screws, ballscrews, epicyclical gears and so on, neglecting driving belts and pulleys.

The motion transmission elements are generally affected by dry friction which may give rise to reversible or irreversible behaviour of the whole system. So, the potential high effect of the dry friction on the dynamic behaviour of the mechanical system requires proper simulation models characterised by high computational accuracy, nevertheless compactness and efficiency.

2 AIMS OF THE WORK

Aims of the work are:

- the study, development, validation and proposal of a general purpose physical and mathematical dry friction dynamic model and, especially, of related simulation computational algorithm.

Contact author: Matteo D. L. Dalla Vedova ¹

¹Department of Mechanical and Aerospace Engineering
Politecnico di Torino
Corso Duca degli Abruzzi 24 – 10129 Torino, Italy.
E-mail: matteo.dallavedova@polito.it

- The problem concerns any type of mechanical device having moving parts, as motion transmissions and gears, affected by dry friction and equipped or not with ends of travel (the authors' specific interest regards the aircraft flight control system actuators); in particular, a new dynamic model/algorithm is proposed in which dry friction is considered to be the sum of both a portion proportional to the load acting on the driven element (through a proper value of efficiency) and a portion independent on it, however being dependent on the relative velocity;
- the detailed description of a general purpose low level language (Fortran, C) numerical algorithm, able to perform the simulations of the dynamic behaviour of the mechanical device being thought in the form of typical flap control system equipped with motion actuators;
- the implementation of a Matlab-Simulink numerical algorithm, having the same purpose as above;
- the simulation of some representative actuation runs in order to validate the proper accuracy of the mechanical device computational algorithm and the analysis of the results.

3 PHYSICAL - MATHEMATICAL MODEL

The dry friction affecting the relative movement of the components of a mechanical system, as, for example, a motion transmission, consists of a force/torque opposing the motion itself, having a value variable as a function of the relative velocity.

In the most of the applications however the relationship between friction force and speed can be represented by the following model (classical Coulomb friction):

- in standstill conditions the friction force can assume any value lower or equal in module to the so said static friction value, opposing the active force and depending on it;
- otherwise the force module has a constant value equal to the so said dynamic friction value, opposing the motion.

This highly nonlinear relationship (discontinuous and undefined in null velocity conditions) gives rise to difficulty in numerical simulation of friction phenomena for the abovementioned purposes. Several different theories have been developed about this problem and the related works fall essentially in two categories concerning the two aspects of the problem itself: continuous and discontinuous types. The aim of the former is mainly the conception of a physical-mathematical model able to represent more and more accurately the relationship between the friction force/torque and the relative velocity of the mechanical parts. Some continuous models consider small elastic displacement (presliding displacement) in the sticking regime and are particularly interesting in the study of specific problems around the null velocity condition (elasto-plastic models [6] et al. [5]), having no further abilities in slipping conditions.

A different problem is considered by the latter category: these models mainly regard the study of mathematical and logical algorithms conceived for the dynamic simulation of mechanical systems characterised by large movements, by integrating, step by step, the equilibrium equations of the moving parts in time dominion: particularly, the algorithms must be able to correctly compute the motion, standstill, stopping, breakaway and reversion conditions.

Typically, in these models the friction force is discontinuous at zero velocity (i.e., in sticking regime) and acts to balance the other forces to maintain zero velocity, if possible. Advantages of discontinuous models are their high performance to simplicity ratio and their wide application field in the classical applied mechanics. However, the conception of the related numerical algorithms is not so simple because their two formulations in conditions of zero and nonzero velocity are completely different; some of the discontinuous friction models most often used are the basic Coulomb (usually implemented by means of a SIGN function), the hyperviscous, the Quinn [4] and Karnopp [3] models, which provide alternative tradeoffs amongst the desirable characteristics of a friction model.

In general, the two categories of problems can be merged together: that is, different friction vs. velocity relationships (nevertheless the Coulomb model) can be suited to simulate every type of motion, standstill, stopping and breakaway algorithm. However, in the most of the applications, considering the friction force/torque vs. velocity relationship, the Coulomb friction approximation is generally satisfactory when the purpose is the dynamic simulation of the behaviour of a multi-component mechanical system characterised by large movements, such as, especially, a motion transmission is, leading a user mechanism and driven by a motor, eventually as a part of a position servo-controller (servomechanism); the load acting on the user mechanism or controlled component may have the same (aiding load) or the opposite (opposing load) sense with respect to the actual motion or, in standstill conditions, to the eventually incipient motion.

In the authors' opinion, the accuracy of the time dynamic simulation of the behaviour of a mechanical system, or position servocontroller, is mainly (nay, much more) dependent on the reliable, robust and trouble-free conception of the algorithm intended to compute motion, standstill, stopping and breakaway conditions; smaller importance is ascribed to the type of friction vs. velocity relationship implemented within the computational algorithm.

Therefore, in the present work, the authors' attention will be much more focused on the former than on the latter considerations. According to these reasons the authors' model employs the Coulomb friction approximation; the consequent relationship among force/torque and velocity (standstill or motion conditions) has been afore described and detailed.

The mathematical and logical algorithm employable to the purpose must be able to perform the behaviour of a movable element affected by friction forces/torques, distinguishing the working condition between the eventual persistence in motion or at a standstill, or the possible motion reversion, breakaway or stopping.

This ability can be important in order to point out some undesired behaviours characterising mechanical devices and particularly servomechanisms.

Like previously said (Coulomb approximation), the evaluation of the dry friction, as a function of the velocity, cannot be described at all by linear models (even though more favourable for the possible analytical solutions of the related dynamic equations). Therefore, whichever attempt of suitable and realistic mathematical modelling requests the use of so complex nonlinearities to advise the employment of numerical computing techniques based on the dynamic simulation in time dominion.

However the techniques of numerical solution mainly employed (and generally reported in nonspecific literature) are based on mathematical models and corresponding computational algorithms that are affected by some shortcomings¹.

The proposed model overcomes the above-mentioned shortages and correctly simulates the behaviour of the mechanical device, as follows:

- selects the correct friction force/torque sign as a function of the relative velocity sense;
- computes the friction force/torque according to the actual external load value acting on the mechanical element;
- distinguishes between aiding and opposing load conditions;
- selects either the static (sticking) condition or the dynamic one (slipping);
- evaluates the eventual stop or breakaway of the previously running or sticking mechanical element respectively;
- keeps correctly in a standstill or motion condition the previously sticking or running mechanical element respectively;
- is able to compute properly the dynamic behaviour of both reversible and irreversible motion transmissions (actuators), taking into account the effects of their eventual mechanical ends of travel.

The proposed model is applied, for demonstration and validation, to the dynamics of an aircraft flight control system and particularly of the actuator-surface assembly, considered as a rigid mechanical element characterised by a single degree of freedom.

¹ Difficulty in implementing the above mentioned friction model in numerical algorithm is rooted in the definition of friction vs. relative velocity v relationship around $v = 0$ and joined computational criteria; in fact, this function is discontinuous with respect to v (and potentially undefined in standstill condition) and the complete definition of the friction value when $v = 0$ is possible by means of the actual external load.

4 FRICTION TORQUE EVALUATION

The dry friction models and computational algorithms available in literature are usually characterized by extremely simplified structures and limited performance; their shortcomings, easily verifiable by means of proper numerical simulations, are particularly emphasized if “integrated” dynamical models are employed, that not only describe the performance of the actuator in Matlab-Simulink taking into account the friction torques but estimating also the possible presence of mechanical ends of travel and their eventual interactions. The algorithm developed by the authors in Matlab-Simulink environment supplies an concrete answer to such problems and, by means of a self-contained subsystem, can describe the effects produced by friction torques on the dynamic behaviour of a generic solid mechanical moving element; the authors’ computational routine can correctly describe many of typical coulomb friction’s effects as well as their interactions with the eventual mechanical ends of travel. The true capabilities of the proposed Coulomb friction computational algorithm are the result of the implementation of a relatively simple but reliable and accurate mathematical model in the versatile Matlab-Simulink environment, so obtaining a self-contained, general-purpose routine employable in a lot of different mechanical applications. In the proposed model, the friction torque is defined in slipping condition as T_{FR} , opposite and invariable with the velocity; T_{FR} is considered as the sum of a component (T_{FR0}) not depending on the load (T_{LD}) and a further one (T_{FRL}) related to the load through a defined value of efficiency. About T_{FRL} , in order to simulate both reversible and irreversible actuators, the proposed model introduces two suitable definitions of efficiency. Usually the efficiency is conventionally defined as the ratio between the output and the input power (dynamic conditions) of a mechanical device as follows:

$$\eta = \frac{T_{OUT} \cdot \omega_{OUT}}{T_{IN} \cdot \omega_{IN}} \quad (1)$$

In (1) it is considered as the input side of the mechanism that in which the torque and angular velocity have the same sense and it is regarded as the output side that in which the torque and angular velocity have the opposite sense.

When the load is opposing the motion, $T_{IN}=T_M$ (motor torque) and $T_{OUT}=T_{LD}$ (aerodynamic load acting on the surface); if the load aids the motion, $T_{OUT}=T_M$ and $T_{IN}=T_{LD}$. If the device is characterised by a constant gear ratio $\tau=\omega_{OUT}/\omega_{IN}$ (typically any mechanism in which the motion is transmitted by the relative movement between conjugate profiles), the efficiency can be intended as the ratio between the output ($\tau \cdot T_{OUT}$) and the input (T_{IN}) torque (related to the same shaft – e.g. the motor one), in any condition:

$$\eta = \frac{T_{OUT}}{T_{IN}} \cdot \frac{\omega_{OUT}}{\omega_{IN}} = \frac{T_{OUT}}{T_{IN}} \cdot \tau = \frac{(\tau \cdot T_{OUT})}{T_{IN}} \quad (2)$$

The efficiency of the mechanism depends on the load condition related to the motion (aiding or opposing). Therefore, if the Coulomb friction model is employed, whatever moving part of a mechanism is characterised, in slipping conditions, by the following two different types of efficiency, which can be intended as follows:

η_{opp} = out/in torque ratio, opposing load

η_{aid} = out/in torque ratio, aiding load

In the opposing conditions the output torque is essentially represented by the load acting on the driven element and the input one by the driving torque (motor or leading element); vice versa in the aiding conditions (in aiding conditions, the usually driven element is considered as the leading one, or input, and the usually driving element as the driven one, or output). By an adequate selection of the values of the above reported efficiencies, it is possible to simulate the behaviour of both reversible and irreversible transmissions. Generally, the efficiencies of the irreversible transmissions are lower than the reversible ones; particularly the aiding efficiency of the irreversible arrangement must be intended as negative.

First of all, the model employs the efficiency value to compute the load dependent friction torque: in fact, the model computes it as the above-mentioned sum of a component not depending on the load (T_{FR0}) and a further one related to the load through the efficiency (T_{FRL}) as:

$$T_{FR} = T_{FR0} + T_{FRL} = T_{FR0} + \left(\frac{1}{\eta_{opp}} - 1 \right) \cdot T_{LD} \quad (3)$$

in opposing conditions and

$$T_{FR} = T_{FR0} + T_{FRL} = T_{FR0} + (1 - \eta_{aid}) \cdot T_{LD} \quad (4)$$

in aiding conditions.

It must be noted that the amount of η_{opp} (defined EtO in the rest of the paper) must lie in the interval between 0 and 1 in order to allow the motor drives the system, while the value of η_{aid} (defined EtA in the rest of the paper) must be not greater than 1; if the mechanical system is reversible the amount of EtA must lie in the interval between 0 and 1, if it is irreversible EtA must be not greater than 0 and the negative value is an expression of the irreversibility degree of the mechanical system. In fact, in slipping and aiding conditions, if EtA=0, the load produces a friction force T_{FR} that is equal and opposing to the load itself ($T_{FR}=T_{LD}$) and their net effect is null, so requiring no action by the motor (neither driving nor breaking); if EtA=-0.5 (or EtA=-1), T_{LD} develops an opposing amount of $T_{FR}=1.5 \cdot T_{LD}$ (or $T_{FR}=2 \cdot T_{LD}$) and their net effect results in a force opposing the motion and amounting to $0.5 \cdot T_{LD}$ (or $1 \cdot T_{LD}$), so requiring a driving action of $0.5 \cdot T_{LD}$ (or $1 \cdot T_{LD}$) by the motor. All these conditions are properly simulated by the algorithm. The model is conceived to compute directly the friction torque T_{FR} in slipping conditions; in case of sticking conditions, the maximum value which can be assumed by friction force is obtained multiplying the slipping one by FSD (static to dynamic friction ratio equal or greater than 1).

When the sticking condition persists, the absolute actual amount of friction force – requested to balance the active force – is not greater than the above considered maximum value. It must be noted that five possible conditions can occur at each computational step:

- Mechanical element initially sticking which must persist in sticking condition, being the absolute value of the active forces (and consequently of friction force) not greater than $FSD \cdot T_{FR}$;
- Mechanical element initially sticking which must breakaway, so turning to slipping condition, being the absolute value of the active forces greater than $FSD \cdot T_{FR}$;
- Mechanical element initially slipping which must keep the slipping condition in the same velocity sense (either when the absolute value of the active forces/torques is greater than T_{FR} , or simply when the element velocity has no-sign reversion within the considered computational step under all the forces acting on it, with respect to its inertia);
- Mechanical element initially slipping which must stop, so turning to sticking condition (having the velocity a potential sign reversion within the computational step, as a consequence of inertia and applied forces);
- Mechanical element initially slipping which must keep the slipping condition, following a motion reversion within the computational step, having the active forces/torques a value greater than $FSD \cdot T_{FR}$ and the sense opposing the initial motion.

The proposed dynamic simulation algorithm is able to distinguish among the conditions a), b), c) and d), solving them within the single considered computational step; the condition e) is performed by means of two following computational steps: the present step is considered as case d) and the following one as case b). It must be noted that this procedure computes a marginally time delayed breakaway (on an average the time-delay amounts to half computing interval). All these abilities are performed in case of both opposing and aiding load with respect to the actual movement or the eventual break-away (incipient motion). It must be noted that, when the external load aids the breakaway, the friction force to be considered is obtained through EtA, while when the external load opposes the breakaway, the incipient motion is performed against a friction force amount depending on EtO; so the procedure, evaluating the eventual breakaway, constrains the T_{FR} amount within two different limits, ruled by EtO and EtA respectively besides by FSD (turning the slipping friction value into the sticking as previously said).

5 ACTUATION SYSTEM MODELLING

5.1 REFERENCE SERVOMECHANISM DESCRIPTION

In order to validate the numerical load depending friction model, its behaviour is studied as a part of a typical electrohydraulic position servomechanism, widely used both in primary and secondary aircraft flight controls; it consists of the following three subsystems:

- a controller subsystem containing a control electronics and a servo-amplifier, typically implementing a PID control law (the present work refers to a pure proportional control law whose behaviour, under friction effects, is much more explicative);
- an electrohydraulic two stage servovalve;
- a hydraulic motor, motion transmission and aerodynamic surface assembly (affected by Coulomb friction), having hard-stops and provided by position transducers closing the control loop.

The full description of the servomechanism employed in the present work and its mathematical model are reported in [1] and [2]. The aforesaid servomechanism belongs to the fly-by-wire paradigm: the pilot's command, depending upon proper transducers, is usually expressed in terms of an electric analog or digital reference signal; this signal is continuously compared via a feedback loop with the actual position of the control surface generating the instantaneous position error as input to the control law. So, the error is processed and transformed into an electric current operating the electrohydraulic servovalve. The servovalve drives an actuator that moves the control surface continuously pursuing, by a proper control law, the reduction of the error between pilot's commanded position and flight surface actual position. The servovalve is a high performance two-stage valve: the corresponding model represents the first stage having a second order dynamics and the second stage as a first order dynamics. The ends of travel of first and second stage are computed. The model of the second stage fluid dynamics takes into account the effects of differential pressure saturations, leakage and variable supply pressure.

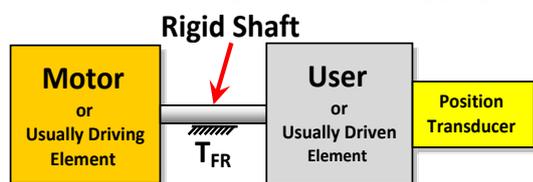


Figure 1 Single mass (1 dof) passive subsystem. Schematic of control system dynamic model.

The actuation system, as shown in Fig. 1, considered in the present paper is composed by a hydraulic motor driving an aerodynamic surface by means of a motion transmission: its model includes inertia, Coulomb and viscous friction and leakage effects through the piston seals developing a not working flow.

5.2 ANALYTICAL MODEL OF SERVOACTUATOR

The position error (Err), coming from the comparison of the instantaneous value of commanded position (Com) with the actual one (XJ), is processed by means of a PID law giving the suitable current input (Cor) acting on the servovalve first stage torque motor; the aforesaid engine torque (expressed as a function of Cor through the torque gain GM), reduced by the feedback effect due to the second stage position (XS), acts on the first stage second order

dynamic model giving the corresponding flapper position (XF) (limited by double translational hard stops).

The above mentioned flapper position causes a consequent spool velocity and, time-integrating, the displacement XS (limited by double translational hard stops $\pm XSM$).

The differential pressure P12 effectively acting on the hydraulic motor, by means of a pressure gain taking into account the saturation effects, is a function of XS and of the total flow through the valve itself.

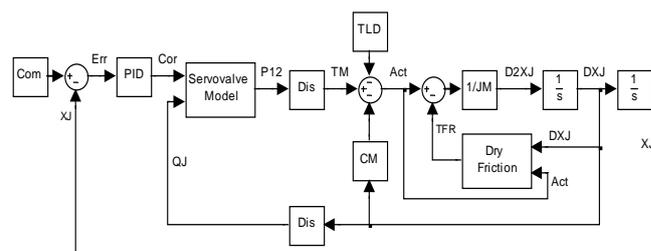


Figure 2 Theoretical actuator dynamics block diagram schematic.

As shown in Fig. 2, the differential pressure P12, through the motor displacement (Dis) and the equivalent total inertia of the surface-motor assembly (JM), taking into account the total load (T_{LD}), the viscous (coefficient CM) and dry friction torque (T_{FR}), gives the assembly acceleration (D2XJ). The equation representing the above mentioned dynamic equilibrium is

$$Act - T_{FR} = J_s \cdot D2XJ \quad (5)$$

where

$$Act = T_M - T_{LD} - C_M \cdot DXJ \quad (6)$$

represents the sum of the active forces/torques, which must be previously known (computed) to evaluate the friction.

The D2XJ time integration gives the velocity (DXJ), affecting the viscous and dry frictions and the linear actuator working flow QJ that, summed to the leakage one, gives the above mentioned pressure losses through the valve passageways. The DXJ velocity time integration gives the actual position (XJ) which returns as a feedback on the command comparison element.

5.3 DYNAMIC EQUATION INTEGRATION

The computational algorithm, originally implemented in FORTRAN environment (as shown in table 1), have been also developed in Matlab-Simulink language (one of the most commonly used languages in engineering applications) and it is shown in Figs. 3, 4 and 5.

Both these algorithms are conceived according to the aforesaid physical friction model and to a general layout not so different from the Karnopp's structure; in fact, both of them are divided in two alternative procedures related to the sticking or slipping condition.

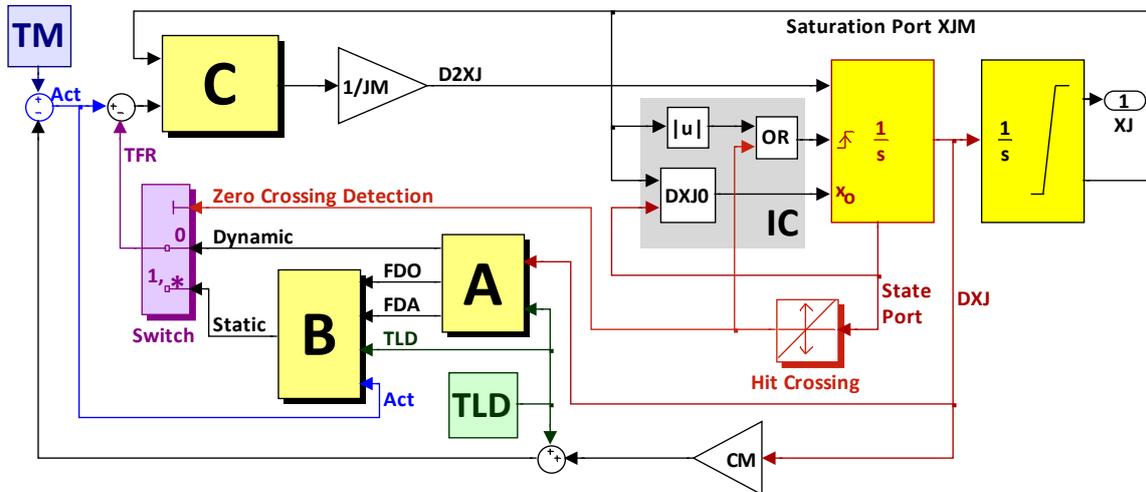


Figure 3 Representation of authors' Matlab- Simulink friction force/torque algorithm.

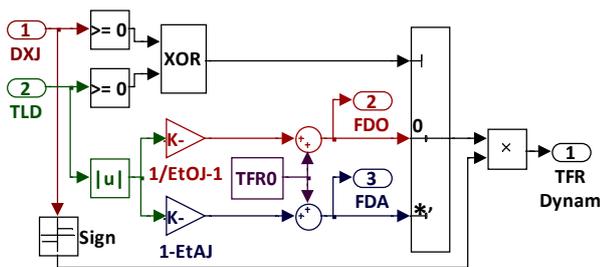


Figure 4 Block A - Representation of dynamic friction force/torque algorithm TFR.

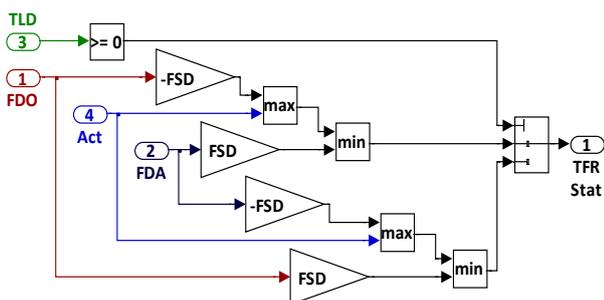


Figure 5 Block A - Representation of static friction force/torque algorithm TFR.

First of all, the algorithm computes the sum Act of the active forces/torques (statement 1 of the computational procedure reported in table 1) and the dynamic values of the friction force/torque (FDO or FDA) as a function of the load T_{LD} in opposing (statement 2) or aiding (statement 3) load run, respectively. The static to dynamic friction ratio

FSD, considered as a constant value in any condition, defines, as a consequence, the static friction values.

In sticking conditions (selected by statement 7), the friction force/torque T_{FR} is considered, by the authors' algorithm, equal to Act and opposing it, but its value must lay within its limits, represented by the static maximum values of friction in eventual opposing ($FSD \cdot FDO$) or aiding ($FSD \cdot FDA$) load breakaway, according to the relative amounts of T_M and T_{LD} (addenda of Act): when T_M is exceeding and opposing T_{LD} , the limit to be assigned to T_{FR} is $FSD \cdot FDO$ (the possible event is the opposing load breakaway), otherwise (T_{LD} exceeding and opposing T_M or aiding it, both exceeding or not) the limit to be considered for T_{FR} is $FSD \cdot FDA$ (possible event: aiding load breakaway), as done in statements 8 or 9 (according to the T_{LD} sense) of the computational routine.

The result is, through the statement 11, an eventual acceleration $D2XJ$ proportional to the eventual excess of Act with respect to the proper T_{FR} limit, having the sense of Act. Synthetically, according to the statements 8 or 9 and 11, the breakaway occurs (in Act sense) only if Act exceeds the proper T_{FR} limit and the excess itself defines slipping condition at the input of the next computational step; otherwise, if Act fails to exceed the T_{FR} limit, no acceleration is performed and the sticking condition persists. In slipping conditions (selected by statement 4), the friction force/torque T_{FR} is considered as the sum of a viscous and a dry (Coulomb, velocity independent) term, opposing the motion; the viscous term is computed, by the coefficient CM, within Act in statement 1, while the dry one is equal to the dynamic value of friction, according to the statements 5 (opposing load run) or 6 (aiding load run).

The result is, by statement 11, an acceleration value $D2XJ$ proportional to the difference between Act and FDO or FDA, having the sense coming from the algebraic difference itself.

By a numerical time integration procedure (as in statement 13, where the simple Euler method is considered), the consequent value of velocity DXJ , characterizing the current step output (considered as input of the following computational step) is computed from the present step input value; the eventual velocity reversion (statements 12 and 14), within the considered computational step (opposite sense between input and output values), must be checked and, if so, the velocity must be imposed equal to zero at the output of the current and so at the input of the next step.

In this way, at the input of the following computational step, the considered mechanical element is necessarily seen in a sticking condition; it seems to be a shortcoming of the algorithm but it is not so. In fact, this measure provides a simple but trouble free method to verify the correct condition (sticking or slipping) to select after a velocity reversion by introducing the computational process into the sticking condition algorithm: so, following the velocity reversion, the sticking condition is maintained if Act is equal or lower than the proper T_{FR} limit or converted into a slipping condition if Act is greater than it. So no specific procedure is necessary for the velocity reversion, having a very small computational error (due to the stop along half computational step, approximately) and no further algorithm burden. The time integration of the velocity DXJ performs the refreshed value of the mechanical element position XJ (statement 15), characterizing the current computational step output (input of the next one), starting from the present step input value. According to the same statement, the mechanical travel is constrained within two hard-stops ($-XJM$, XJM); when a hard-stop is reached, the impact is computed as quite inelastic and, as a consequence, the velocity DXJ is set equal to zero at the output of the current and so at the input of the next step (statements 16 and 17). The same statements allow the eventually correct departure from the considered hard-stop, only preventing negative or positive velocity values when the upper (negative, statement 16) or lower (positive, statement 17) hard-stop is engaged, respectively; the purpose of this computational structure is the prevention of eventually delayed departure from the hard-stop, caused by possible very small computed acceleration values $D2XJ$, in particular conditions.

PROPOSED SIMULINK ALGORITHM: as regards the authors' Simulink algorithm, it is able to simulate the dynamic behaviour of a second order mechanical system computing the proper static/dynamic value of the friction force by means of the blocks A and B (Fig. 4 and 5 respectively) and taking into account the eventual effects of hard-stops. As shown in Fig. 4, the block A (according to statements 2, 3, 5 and 6 of table 1) implements the equations (3) and (4), calculating the dynamic value of the friction force/torque (FDO or FDA) as a function of load

T_{LD} and velocity DXJ ; in particular, such algorithm is able to select between opposing and aiding conditions (employing FDO or FDA by means of the Switch block) and evaluate the correct sign of the so obtained friction force/torque T_{FR} (taking into account the DXJ sign by means of a block implementing the signum function).

Starting from the abovementioned dynamic values of the friction force/torque, the block B (shown in Fig. 5 and according to the statements 8 and 9) computes the related static friction force/torque T_{FR} equal and opposite to Act, but limiting it within the corresponding static maximum values of friction in opposing (FSD·FDO) or aiding (FSD·FDA) conditions.

As regards the breakaway detection (statements 4), the authors' Simulink algorithm implements this routine by means of a switch block that, as a function of instantaneous value of DXJ (coming from the integrator state port), selects among sticking and slipping condition (by means of a hit crossing block) and, so, gives in output the proper value of static or dynamic friction force T_{FR} (Fig. 3).

As shown in Fig. 3, the constrain of the mechanical travel within the hard-stops ($-XJM$, XJM) is achieved limiting opportunely the output of the second integrator (according to statement 15).

The velocity resets, due to the eventual DXJ reversion or engagement of a hard-stop, are performed by means of the joint action of the block IC and C (Fig. 3). The block IC, detecting the DXJ reversion (by means of the hit crossing block) or the hard-stop reaching (detected by means of the saturation port of the integrator), is able to reset the output of the corresponding integrator (statement 14) and, if $XJ=|XJM|$, assign the proper reset DXJ value (statements 16 and 17). The block C, simulating the effect of the hard-stop reaction force, nullify the $D2XJ$ value until the controlled element is pushed against the hard-stop (avoiding undesired integrator windup).

6 SIMULATION RESULTS

In general, as it is well known, when a load is applied to the output side of an irreversible mechanical system, in sticking conditions it develops a friction force/torque having a value potentially exceeding the load itself (actually equal to it, if no driving torque is present on the input side), so preventing any incipient movement (if only the load is applied); otherwise, in slipping conditions, the system develops a friction force/torque actually exceeding the load itself, so progressively stopping the current motion, if no action is applied on the input side of the mechanical system. In sticking conditions, the breakaway is possible or in slipping the motion may persist only as a consequence of a proper driving action applied on the input side. On the contrary, as the reader grasps immediately, if the motion transmission is reversible the load alone, applied on the output side, may be capable of performing the system breakaway in sticking or the motion persistence in slipping conditions.

In the light of these considerations, it is possible to examine closely the simulations results, as follows.

In Figs. 6, 7, 8, 9 and 10 the position command Com is given in form of a step change at time = 0.1 s and a step load T_{LD} is applied from 0 to 74500 N·m at time = 0 s opposing (Figs. 6 and 8) or aiding (Figs. 7, 9 and 10) the commanded actuation rate, respectively. Further, the nominal differential value between supply and return pressures PSR, provided to the servomechanism by the hydraulic system, is set at 26 MPa. This pressure really acts as a constant value at the beginning and at the end of every simulation, whereas, in the central portion of it, a temporary marked pressure drop from 26 to 5 MPa and following recovery to 26 MPa is considered as a hydraulic system malfunctioning, affecting the actuation travel: it is assumed as a continuous time function, having a linear decrease from 26 to 5 MPa since time = 0.3 s till to 0.4 s, followed by 5 MPa constant value and then a linear increase restoring 26 MPa, as a final value, within the time interval 0.6 to 0.8 s.

In Figs. 6 and 8, the position command is given from an initial null position to 1.5 degs.

In Figs. 6 and 7 a reversible mechanical subsystem is assigned to the considered servomechanism, which develops a movement following the step load application, reaching a position error value able to balance the load itself, by means of the servomechanism motor action (produced by its position stiffness) together with the friction torque T_{FR} .

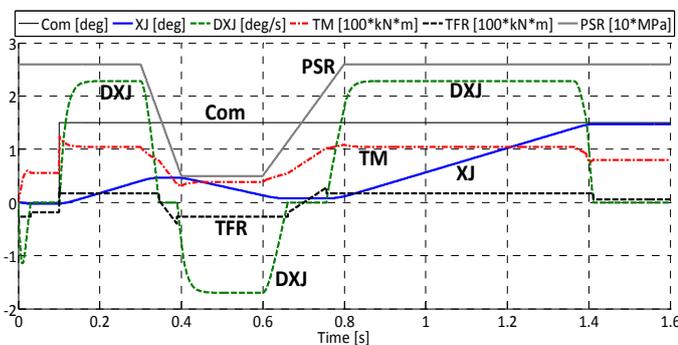


Figure 6 Reversible mechanical subsystem deployment.

In Fig. 6, when the position step command Com is applied, the large error value produces the quick breakaway of the system, overcoming both friction torque and load, consequently reaching a constant actuation rate condition; in it, the action of the hydraulic system differential pressure balances the pressure losses through the servovalve passageways, the viscous friction, the load dependent and independent dry friction and the external load and it is related to the values characterising the motor displacement, the servovalve ends of travel, pressure and flow gains and so on. When the supply pressure drop starts, the actuation rate decreases till to a full system stop.

This sticking condition persists till to the undesired system breakaway developing an actuation rate sense opposing the input command (load exceeding the sum of driving and friction torques); in this condition the authority is acquired by the external load, able to produce an “aiding” movement, overcoming the position command, as a consequence of the motion transmission reversibility, coupled with the very low supply pressure, as it can be expected. When the supply pressure recovery starts, the undesired “aiding” movement decreases till to a new sticking condition, followed by the desired breakaway towards the commanded position. It occurs when the supply pressure is able to produce a driving torque amount exceeding the sum of load and friction torque, so restoring the correct load opposing actuation rate.

All the sticking conditions are characterised by an absolute value of the difference between driving torque T_M and load T_{LD} lower than the maximum static friction torque (opposing if T_M overcomes T_{FR} , aiding if vice versa). At time = 1.41 s the commanded position is reached, having a final “error” caused both by the load and the friction torque.

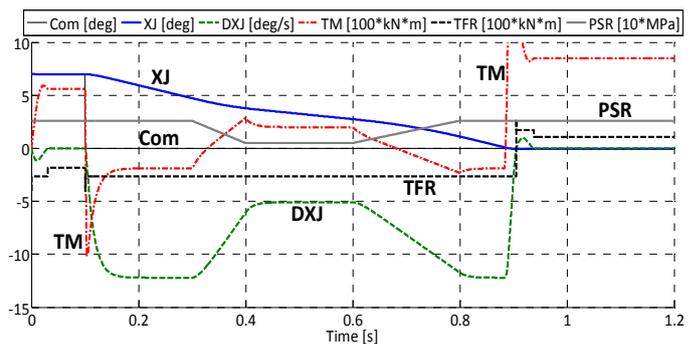


Figure 7 Reversible mechanical subsystem retraction.

In Fig. 7 the position command is given as a step change from an initial 7 degs position to 0 degs, so producing an actuation rate sense characterised by aiding load. Following the position step command application, a constant actuation rate condition is reached; in it, hydraulic system differential pressure PSR and external load balance servovalve passageways pressure losses, viscous friction, load dependent and independent dry friction. When the supply pressure drops, the actuation rate decreases, but no sticking conditions is reached, as a consequence of aiding load and motion transmission reversibility, involving friction torques lower than load actions. When the supply pressure restores the actuation rate increases again till to the previously seen value, so reaching the commanded position at time = 0.94 s. Figures 8, 9 and 10 regard an irreversible mechanical subsystem, assigned to the considered servomechanism which is not interested by any movement following the step load application, having the sole load dependent friction torque the ability to balance the load itself, without servomechanism motor action.

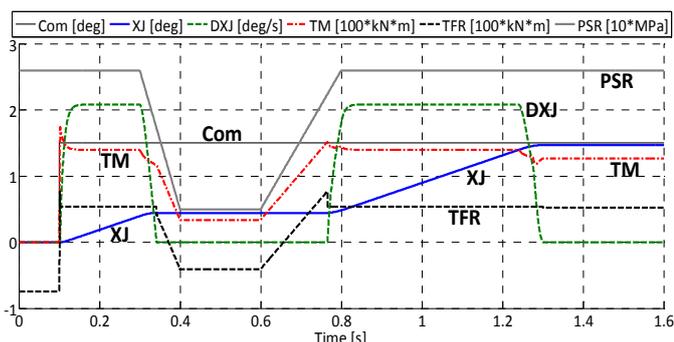


Figure 8 Irreversible mechanical subsystem deployment low “level” of irreversibility ($\eta_{aid} = -0.2$).

In Fig. 8, when the position step command is applied, following a quick acceleration transient, a constant actuation rate condition is reached; in it, the torques balance is quite similar to the case reported in Fig. 6.

When the supply pressure drops, the actuation rate decreases till to a sticking condition, without any back movement aided by the load, because of the mechanical system irreversibility; in fact, the inability of the motor element to balance the load has no effect on the system, because, whatever value the load assumes, the friction torque, produced by the load itself and opposing it, can be eventually greater than it and actually equal to it, so preventing any movement not induced by input side actions. The recovery of the supply pressure performs a new load opposing system breakaway towards the commanded position. It occurs for the same reasons reported in the case of Fig. 6. At time = 1.3 s the commanded position is reached, having a final “error” related both to the load and the friction torque.

In Figs. 9 and 10, the position command Com is given from an initial 4 degs position to 0 degs. Following the position step command application and the consequent quick acceleration transient, a constant actuation rate condition is reached; in it, the torques balance is quite similar to the case considered in Fig. 7.

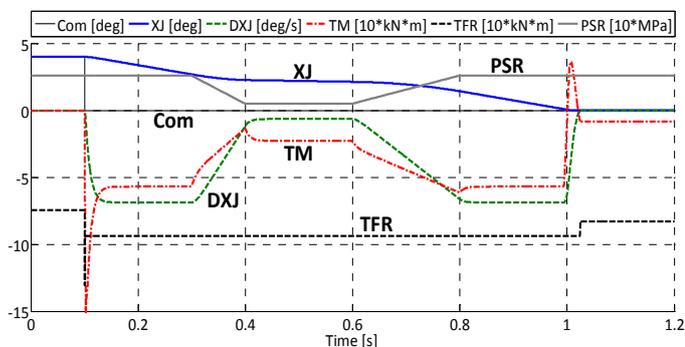


Figure 9 Irreversible mechanical subsystem retraction low “level” of irreversibility ($\eta_{aid} = -0.2$).

In Fig. 9, when the supply pressure drops, the actuation rate decreases, but no sticking condition is reached, as a

consequence of the aiding load, notwithstanding the motion transmission irreversibility; in fact, in this condition, the friction torque is greater than the load, but their net effect is lower than driving torque produced by the differential supply-return pressure PSR so allowing the development of an actuation rate as commanded. It must be noted that the torque related to the pressure losses within the valve passageways produced by the actuation rate, balances the abovementioned torques acting on the mechanical system. This behaviour is related to the low “level” of irreversibility, selected for the present motion transmission ($\eta_{aid} = -0.2$). When the supply pressure restores, the actuation rate increases again till to its initial value, so reaching the commanded position at time = 1.02 s.

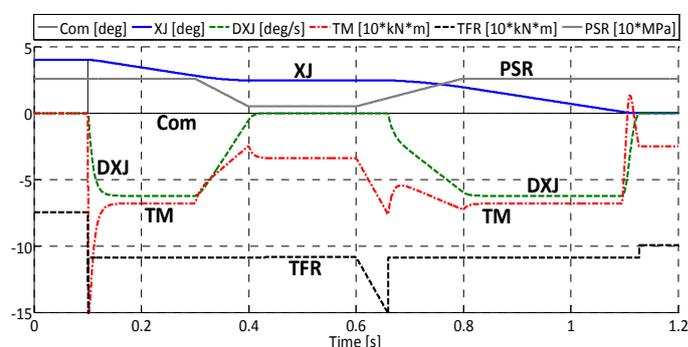


Figure 10 Irreversible mechanical subsystem retraction higher “level” of irreversibility ($\eta_{aid} = -0.4$).

Fig. 10 is quite similar to Fig. 9 apart from the higher “level” of irreversibility of the motion transmission ($\eta_{aid} = -0.4$). In this case, when the supply-return pressure drops, the actuation rate decreases reaching a sticking condition as a consequence of the mechanical irreversibility, notwithstanding the aiding load; in fact, the higher excess of friction torque with respect to the load overcomes the driving torque related to the differential supply-return pressure PSR, so performing a motor stall condition. When PSR restores, the system performs a new breakaway developing the initial actuation rate value and, then, reaching the commanded position at time = 1.28 s. The comparison between Figs. 9 and 10 proves the ability of the algorithm to take correctly into account the effects of the irreversibility “level”, so defining the ratio between friction torque and load in any condition.

In Figs. 11, 12, 13, 14, 15, 16 and 17 the differential supply-return pressure PSR provided to the servomechanism by the hydraulic system is assumed equal to its 26 MPa nominal value along the whole simulation time.

In Figs. 11, 12 and 13 a step load T_{LD} is applied from 0 to 74500 N·m at time = 0.02 s, restoring 0 N·m at 0.2 s by a second step change; the actual initial position is set at 0.1 degs and the related command keeps the same value till to time = 0.07 s; then, a position ramp command follows from 0.1 degs, at time = 0.07 s, characterised by -0.5 degs/s slope.

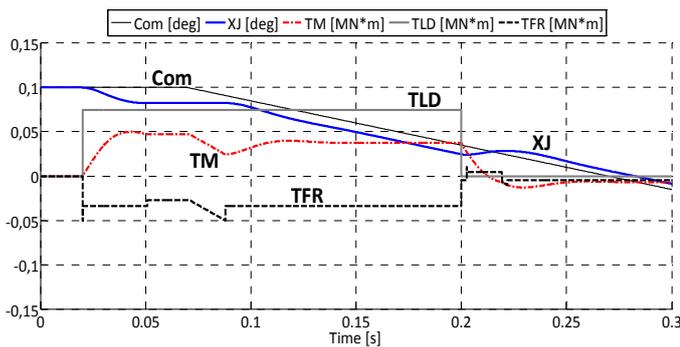


Figure 11 Reversible mechanical subsystem retraction PSR nominal value (26 MPa) and T_{LD} step load.

Fig. 11 regards a servomechanism equipped with reversible mechanical subsystem; it develops, at time = 0.02 s, a movement following the first step load application reaching a position error value able to balance the load itself by the motor action and the friction torque.

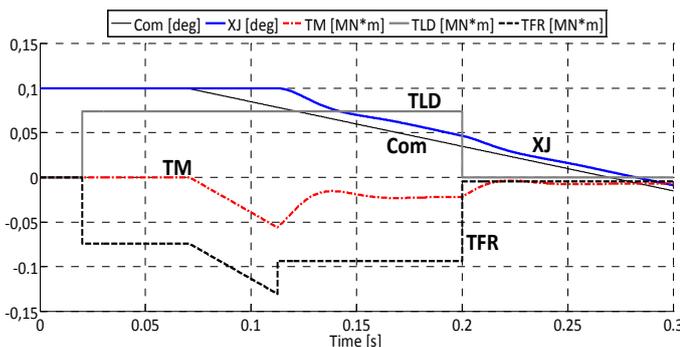


Figure 12 Irreversible mechanical subsystem retraction PSR nominal value (26 MPa) and T_{LD} step load.

Fig. 12 shows the irreversible mechanical subsystem behaviour, performing no movement in the same step load conditions, as it can be expected. When the position ramp command Com starts, the system breakaway occurs following a defined time delay (resolution), due to the small, but variable, initial error value together with the static friction torque; the system moves when the driving torque, related to the abovementioned position error, overcomes the net torque between static friction and load.

It must be noted that, having the ramp command a negative slope and the servomechanism purely proportional control law, the reversible (Fig. 11) or irreversible (Fig. 12) system breakaway occurs starting from a positive (Fig. 11) or negative (Fig. 12) value of position error. In no load conditions the sign of the position error must be the same of the ramp command slope (ComR): if the net effect between load and torques connected to viscous effects, friction and pressure losses across the valve (small actuation rate) aids the commanded movement (situation necessarily connected to an aiding load acting on a reversible mechanical system),

the slope and position error signs are opposite (as shown in Fig. 11 till to time = 0.2 s).

The system irreversibility (as in Fig. 12) produces the opposite behaviour, as a consequence of the excess of the friction torque with respect to the load. As previously reported, when the load value returns to 0 N·m (time \geq 0.2 s), both the reversible and irreversible systems develop position error having the same sign of ComR, as it can be expected and it is correctly shown in Figs. 11 and 12.

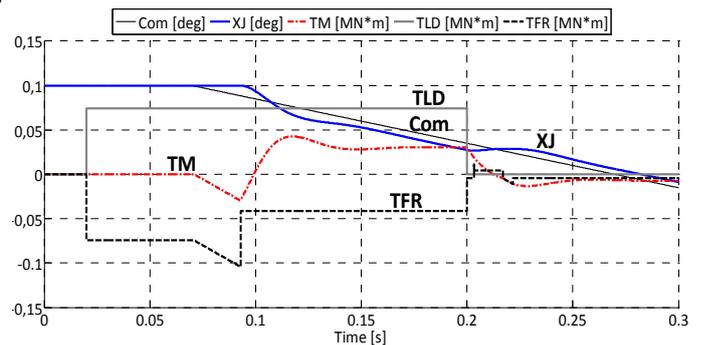


Figure 13 Statically irreversible / dynamically reversible mechanical subsystem retraction

Fig. 13 is referred to a particular but possible behaviour of the mechanical subsystem: the efficiency in dynamic aiding condition is defined as positive (dynamically reversible system), but the ratio between static and dynamic friction torque is assumed sufficiently high to perform a static friction torque exceeding the corresponding load (while the dynamic one is lower). In fact, as clearly shown in Fig. 13, the position error in breakaway condition is proper of an irreversible system, but, in steady state, the behaviour is characteristic of a reversible one.

It must be noted that, also in this case, the authors' computational algorithm is quite able to correctly reproduce the expected behaviour: in fact, in aiding conditions, the average position error sense concerning the reversible system is the opposite of the required actuation rate, because of its positive value of efficiency, involving friction torques (opposing) lower than the related load ones (aiding), thus producing an aiding net torque, requiring a braking action to the hydraulic motor. Therefore, the position error must oppose the commanded movement if the aiding net torque exceeds the braking action represented by the viscous damping of the mechanical subsystem and mainly by the pressure losses through the valve passageways related to the flow required by the motor speed. Instead, in case of irreversible system, the average position error sense is the same of the required actuation rate, because of its negative value of efficiency, involving friction torques (opposing) exceeding the related load ones (aiding), thus producing an opposing net torque, requiring a driving action to the hydraulic motor; therefore, the position error must aid the commanded movement.

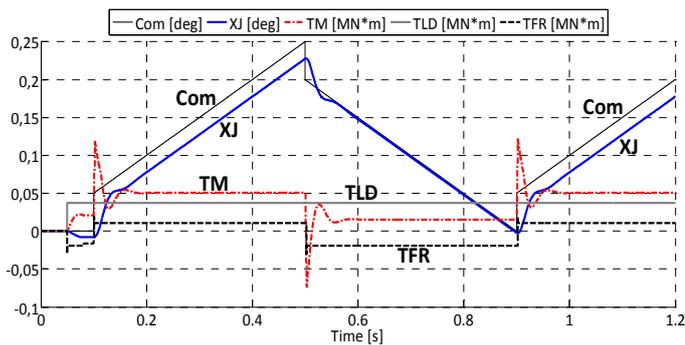


Figure 14 Reversible mechanical subsystem actuation.

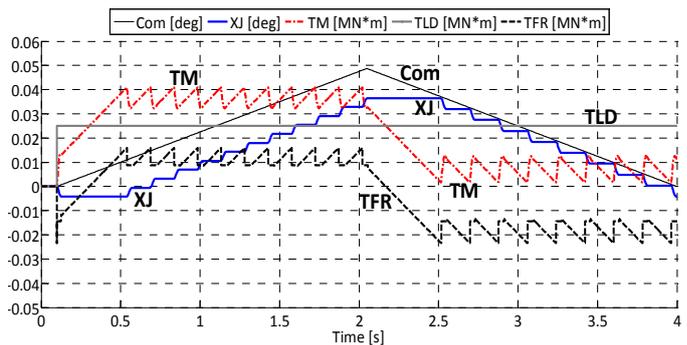


Figure 16 Reversible mechanical subsystem actuation.

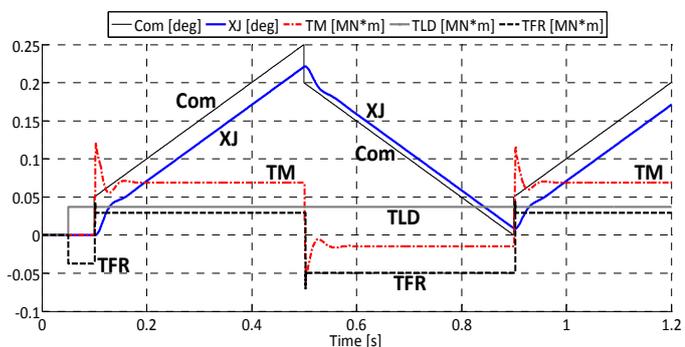


Figure 15 Irreversible mechanical subsystem actuation.

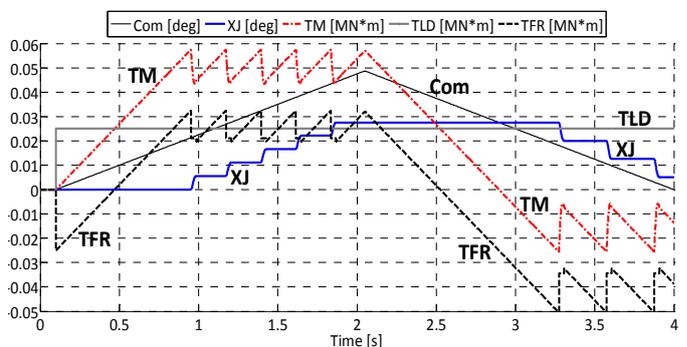


Figure 17 Irreversible mechanical subsystem actuation.

In Figs. 14 and 15 a step load is applied from 0 to 37243 N·m at time = 0.1 s, keeping it constant till to the simulation end; the actual initial position is set at 0 degs and the same value is given to the command, constant till to time = 0.1 s, in which time a position step command of 0.05 degs is applied. Then a position ramp command follows from 0.05 degs at time = 0.1 s having a growth rate of 0.5 degs/s till to 0.5 s. At time = 0.5 s a step command of -0.05 degs is applied, the followed by a decrease of -0.5 degs/s till to 0.9 s. Successively, the sequence from 0.1 s to 0.9 s is repeated. The step position commands applied at the beginning of each ramp has the purpose of producing a brief transient to the steady-state ramp response and, so, removing the behaviour difference between reversible and irreversible system in break-away condition.

The comparison between Figs. 14 and 15 shows that the delay between commanded and actual position is lower in case of a reversible than irreversible system (eventually opposite to the commanded actuation rate when the load is aiding); the detailed view in Fig. 14 puts in evidence that the actual position leads the commanded one in condition of steady-state low rate ramp command when the system is reversible and the load aids the motion. It must be noted that the examined situation is the consequence of properly low actuation rate (but higher enough to prevent stick-slip), sufficiently high value of aiding load T_{LD} and low value of load independent friction force.

In Figs. 16 and 17 a step load is applied from 0 to 25000 N·m at time = 0.1 s, keeping it constant till to the simulation end; the actual initial position is set at 0 degs and the same value is given to the command, constant till to time = 0.1 s. Then a position ramp command follows from 0 degs at time = 0.1 s having a growth rate of 0.025 degs/s till to 2.05 s, followed by a decrease of -0.025 degs/s. In these cases the required actuation rate is sufficiently low to perform a stepping dynamic response (stick-slip) as a consequence of the difference between static and dynamic friction forces.

In general, the operative conditions having a lower value of efficiency are characterized by higher steps and (as a consequence of a defined input ramp slope) longer period. In fact, low efficiency involves large values of friction force/torque related to a given external load, so performing high difference between its static and dynamic values, for a defined FSD (static/dynamic) ratio.

According to these considerations, as it can be observed in Figs. 16 and 17, having the same data set, the irreversible system is characterised by higher steps than the reversible one; further, the aiding load travel performs higher steps than the opposing one.

Consequently, the comparison between Figs. 16 and 17 puts in evidence higher values of average position error regarding the irreversible system than the reversible one along the actuation travel in opposing conditions, as a result of its lower efficiency.

7 CONCLUSIONS

The simulations show the proper accuracy of the proposed algorithm taking into account the effects of the dry friction and of the ends of travel on the behaviour of the actuators. It must be noted the ability of the proposed model to describe correctly the dynamic/static behaviour of both reversible and irreversible types, employing the proper values of the respective efficiencies.

So, the algorithm developed by the authors supplies an effective answer to the necessity of accurate tools in evaluating the effects produced by friction forces or torques and ends of travel acting on a generic mechanical device, as an actuator, by means of a self-contained Simulink computational routine.

8 TABLE 1:

FORTRAN Listing of the Authors' Friction Algorithm

```

01 - Act =TM-TLD-CM-DXJ
02 - FDO=TFR0+(1/EtO-1)*ABS(TLD)
03 - FDA=TFR0+(1-EtA)*ABS(TLD)
04 - IF(DXJ.NE.0.) THEN
05 -   TFR=SIGN(FDO,DXJ)
06 -   IF(DXJ.FR.LT.0.) TFR=SIGN(FDA,DXJ)
07 - ELSE
08 -   TFR=MIN(MAX(-FSD-FDA,Act),FSD-FDO)
09 -   IF(TLD.LT.0.) TFR=MIN(MAX(-FSD-FDO,Act),FSD-FDA)
10 - ENDIF
11 - D2XJ = (Act-TFR)/JM
12 - Old = DXJ
13 - DXJ = DXJ+D2XJ*DT
14 - IF (Old-DXJ.LT.0) DXJ = 0
15 - XJ=MIN(MAX(-XJM, XJ+(Old+DXJ)*DT/2),XJM)
16 - IF(XJ.EQ.-XJM) DXJ=MAX(0.,DXJ)
17 - IF(XJ.EQ.XJM) DXJ=MIN(DXJ,0.)

```

9 LIST OF SYMBOLS

Act sum of the active torques
CM viscous damping coefficient of the system
Com input command
D2XJ controlled element angular acceleration
DXJ controlled element angular rate
FSD static to dynamic friction ratio
JM moment of inertia of the system
PSR differential supply-return pressure
T_{LD} aerodynamic load acting on the surface
T_{FR} friction torque
T_{FRO} load independent friction torque component
T_{FRL} load dependent friction torque component
T_M motor torque
XF servovalve first stage flapper position
XJ controlled element angular position
XS servovalve second stage spool position
 η general out/in torque ratio
 η_{opp} out/in torque ratio, opposing load
 η_{aid} out/in torque ratio, aiding load

10 LIST OF SYMBOLS USED IN TABLE 1

EtA = η_{aid}
 EtO = η_{opp}
 FDA dynamic friction torque, aiding load
 FDO dynamic friction torque, opposing load
 TFR = T_{FR}
 TFR0 = T_{FRO}
 TLD = T_{LD}
 TM = T_M

11 LITERATURE

- [1] Borello L., Villero G., Flap control system actuators: mathematical and computational model for dynamic simulation, *European Congress on Computational Methods in Applied Sciences and Engineering ECCOMAS 2000*, Barcellona, Spain, 11-14 September 2000.
- [2] Borello L., Dalla Vedova M. Load dependent coulomb friction: a mathematical and computational model for dynamic simulation in mechanical and aeronautical fields. *International Journal of Mechanics and Control (JoMaC)*, Vol. 07, No. 01, pp. 19-30, 2006.
- [3] Karnopp D., Computer simulation of stick-slip friction in mechanical dynamic systems, *Journal of Dynamic Systems, Measurement, and Control*, Vol. 107, pp. 100-103, March 1985.
- [4] Quinn D. D., A new regularization of Coulomb friction," *Trans. Of ASME: Journal of Vibration and Acoustics*, vol. 126, no. 3, pp. 391-397, 2004.
- [5] Kikuuwe R., Takesue N., Sano A., Mochiyama H., Fujimoto H., Fixed-step friction simulation: from classical coulomb model to modern continuous models, *International Conference on Intelligent Robots and Systems IEEE/RSJ 2005*, pp. 3910-3917, June 2005.
- [6] Dupont, P., Armstrong, B., Hayward, V., Elasto-plastic friction model: contact compliance and stiction, *Proceedings of the American Control Conference*, vol.2, pp. 1072 - 1077, June 2000.

DESIGN PROCEDURE OF A BREECH MECHANISM WITH DOUBLE HELIX SHAFT

Cristiana Delprete*

Elvio Bonisoli**

Daniele Tomasi***

* Politecnico di Torino, Mechanical and Aerospace Eng. Dept., C.so Duca degli Abruzzi, 24 - 10129 Torino ITALY

** Politecnico di Torino, Management and Production Eng. Dept., C.so Duca degli Abruzzi, 24 - 10129 Torino ITALY

*** Corso Machiavelli, 119 - 10078 Venaria Reale (TO) ITALY

ABSTRACT

The breech mechanism with double helix shaft is an innovative system for breech blocking that provides a mechanical support to firearms fed with caseless ammunition. In a small firearm that has high rate of fire, the breech mechanism failure occurs mainly due to the fatigue phenomenon. In this paper it is described the design procedure of the core parts of the mechanism and the fatigue verification is reported.

Keywords: breech mechanism, caseless ammunition, small arms, mechanical design, fatigue verification.

1. INTRODUCTION

In small arms technology, one of the latest innovations is the use of caseless ammunition [1-3]. The use of this kind of ammunition allows the substantial reduction of weight, increasing the cartridges transport capability of the infantry operators, the reduction of logistic costs of the army, the reduction of production costs and the reduction of the environment pollution generated by the discharge of the brass case of the traditional ammunition [4-6].

By counterparts this new type of cartridges requires the solution of various technical requirements, such as:

1. Need of a solid propellant that keeps the projectile and the primer together with the propellant itself substituting the case;
2. Need of a thermally stable propellant that is exposed to the hot wall of the chamber, in order to avoid the auto cook-off phenomenon;
3. Need of sealing on the breech bore previously done by the case radially deformed under the inner pressure until the sealing over the inner wall of the barrel.

The first task can be solved, for instance, using a moulded propellant charge mixed with a resin that improves the mechanical properties of the pressed powder [7].

The second problem can be reduced with the use of thermally stable mixtures of propellant and with the use of a coating that reduces the sensibility of the charge to the temperature [7, 8].

The last point is the most important in order to get an high efficiency of the gun system, obtaining the maximum conversion of energy from the gas pressure generated by the combustion into the kinetic energy of the projectile.

In a firearm, the breech mechanism is the device responsible to provide the closure of the breech bore of the barrel.

Various types of breech mechanisms have been developed through the years, one of the last is the Breech Mechanism with Double Helix Shaft (BMDHS) proposed in 2011 [9]. This application aims to solve the problem of the breech sealing in firearms that fires caseless ammunition with a conical sealing that is pushed on the breech bore.

In traditional weapons, the sealing due to the deformed case is not affected by the clearances of the breech mechanism, but in weapons that fire caseless ammunition clearances reduce the efficiency, especially in long arms such as rifles or carbines where the performances of the caseless guns reduced by the gas leakage that can influence the positioning of this new technology on the market.

Therefore it is important the use of a mechanism that offers the perfect sealing of the breech bore.

Contact author: C. Delprete¹, E. Bonisoli², D. Tomasi³

E-mail: cristiana.delprete@polito.it
elvio.bonisoli@polito.it
daniele.roberto.tomasi@gmail.com

2. DESCRIPTION OF THE MECHANISM

The BMDHS aims to maintain the direction of the movements of the breech mechanism parallel to the barrel axis in order to simplify the automation of the mechanism

and to obtain low operation time with consequent high ratio of fire. It aims also to provide the sealing for the gas at the breech bore, which requires great forces, assuring enough manoeuvring space for the cartridge, while requiring a low

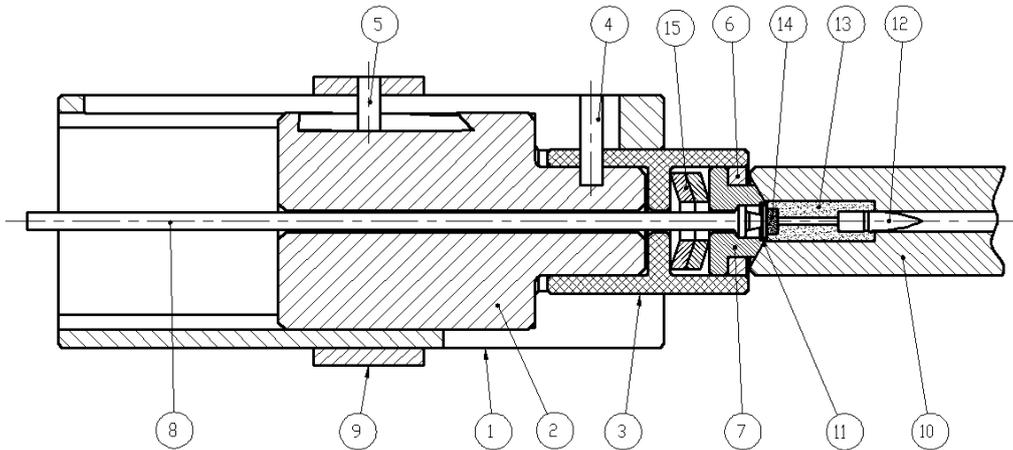


Figure 1 Sketch of the BMDHS section: 1 housing, 2 shaped shaft, 3 sleeve, 4 actuation pin, 5 command pin, 6 locking ring nut, 7 breechblock, 8 firing pin, 9 cursor, 10 barrel, 11 elastic ring, 12 projectile, 13 propellant, 14 primer, 15 Belleville springs.

input force because it must be manually operated to load the first shot. Furthermore the mechanism has to be automatic, fast and compact.

In order to satisfy all of these requirements, the BMDHS uses a conical sealing, a force reduction and a variation of constraints configuration to grant respectively the gas sealing, the manual operability, quickness and compactness.

Moreover the BMDHS motion occurs along the barrel axis, providing the easy automation of the mechanism.

2.1 PARTS DESCRIPTION

Figure 1 represents a sectional view of the mechanism and the BMDHS parts are briefly described [10].

The sleeve 3 (Figure 2) is free to slide inside the forward hole of the housing 1. The actuation pin 4 is coupled with the forward helical groove on the shaped shaft 2 (Figure 3) and it is bounded to move horizontally inside the groove in the housing 1.



Figure 2 Sectional view of the sleeve.



Figure 3 Rendering of the shaped shaft.

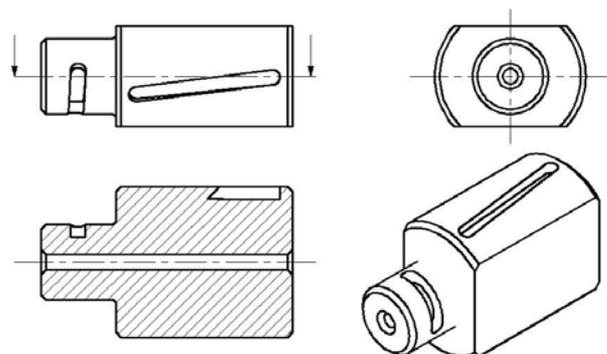


Figure 4 Shaped shaft views.

The command pin 5 is bounded to move horizontally inside the groove of the housing 1 and engages also with the rear helical groove of the shaped shaft 2. The cursor 9 is

bounded to move axially along the housing 1 and it is the element that drives the mechanism.

The cursor is operated during the forward stroke by a rectangular compression coil spring with circular wire and low stiffness while, during the rearward stroke, according to the type of the firearm, the cursor can be operated rearward by a gas recovery device or manually. The shaped shaft 2 is housed and relies radially on the circular sector of the shaped hole inside the housing 1.

2.1.1 Shaped shaft

The shaped shaft (Figures 3 and 4) is composed by an anterior cylindrical side, on which a helical groove is milled, and by a rear cylindrical side, with higher diameter, on which another helical groove is milled. The rear cylindrical side has also two flat parallel surfaces. The shaft has a hole inside to house the firing pin.

2.1.2 Housing

The housing (Figure 5) is composed by a single parallelepiped block with rectangular base, developed along the axial direction of the mechanism. Inside the housing there is a shaped hole with two circular sectors and two flat parallel surfaces that reproduce the shape of the shaped shaft.

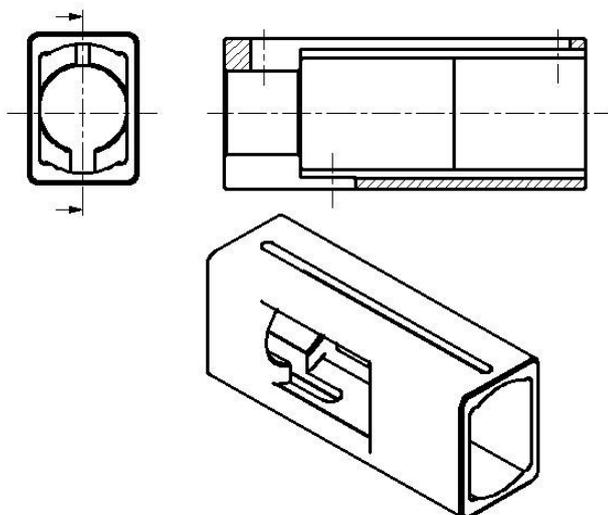


Figure 5 Housing views.

On the front surface there is a hole inside which slides the sleeve when this is in forward position. The sides' walls have an opening in the front side of the housing, where the shaped shaft is during the locking stroke.

On the top surface there is a passing through linear groove that is parallel to the longitudinal axis of the housing. On the bottom surface, in the front side there is a groove that allows to the ammunition the manoeuvring space needed to exit from the magazine and enter inside the barrel.

2.2 DESCRIPTION OF THE OPERATION

The mechanism movement is ideally split in two strokes: a rapid stroke and a locking stroke.

During the rapid stroke, the sleeve performs a long stroke, long enough to assure sufficient manoeuvring space to the ammunition. During the locking stroke the sleeve performs a very short stroke, long enough to force the breechblock on the barrel breech bore with enough strength to ensure the sealing under the action of the gas pressure.

The shaped shaft, during the rapid stroke, performs a linear axial movement, whilst, during the locking stroke, the shaped shaft performs a rotation around its longitudinal axis. The kinematics of the shaped shaft depends on the housing that imposes different boundaries in relation to the position assumed by the shaped shaft and by the cursor that imposes its speed to the mechanism.

During the rapid stroke the housing bounds the shaped shaft to be coaxial via the cylindrical surfaces and to not rotate around its axis, through the flat parallel surfaces.

During the locking stroke, the housing bounds the shaped shaft to be coaxial through the cylindrical surfaces and leaves it free to rotate around his axis thanks to the windows opened on the side walls of the housing. During this stroke the breechblock is in contact with the barrel breech bore and this last one gives the constrain reaction to the breechblock which is equal to the locking force. Therefore the shaped shaft will tend to rely on the rear planar surfaces of the side windows of the housing.

During the rapid stroke, the kinematic relation between the cursor and the sleeve is a rigid linear translational movement or rather the speed of the sleeve equals the cursor speed, thus there is a unitary transmission ratio. In facts, the shaped shaft is bounded to not rotate thus there is no relative motion between command pin (and the cursor) and its respective helical groove of the shaped shaft. Similarly there is no relative motion between the actuation pin (or the sleeve) and the respective helical groove of the shaped shaft, therefore cursor, command pin, shaped shaft and actuation pin move linearly rigidly.

During the locking stroke the kinematic relation between the cursor and the sleeve is given by the wrap angles of the two helical grooves of the shaped shaft and they are chosen in order to get an opportune transmission ratio to realize a reduction of the force that the coil spring must operate on the cursor with respect to the sealing force of the breechblock. In facts, the forward movement of the cursor is equal to the movement of the command pin and this last one engaging the respective helical groove on the shaped shaft make the shaft rotate around its axis and this rotation is transmitted rigidly to the other helical groove of the shaped shaft which engages the actuation pin that moves forward. Therefore the wrap angles of the two helixes must agree. The transmission ratio between the command pin and the actuation pin is a function of the wrap angles and radii of the two helixes. Choosing opportune radii and wrap angles, it is possible to obtain great force reduction from the sealing force to the operating force, obtaining as counterpart a high reduction of the stroke from the cursor to the sleeve. For this reason the mechanism has two operating modes: the locking stroke requires a high stroke

at the cursor, therefore it is limited to cover the sleeve displacement until the Belleville springs have enough deformation to press the breechblock with the required sealing force, whilst the rapid stroke allows to the breechblock to move rearward enough to let the ammunition enter into the barrel requiring to the cursor an additional stroke that equals the stroke of the breechblock.

2.3 DYNAMICS OF THE MECHANISM

As a first step in the fatigue life verification, the interaction of the mechanism has to be studied, in order to determine the intensity of the forces that stress the components.

2.3.1 Rapid stroke configuration

In this configuration the boundaries prevent the shaft from rotation furthermore the command pin can only move along the groove into the housing. Then the slide of the command pin results in a rigid axial translation of the assembly composed by the command pin, the shaft and the actuation pin. Therefore the force acting on the command pin is equal to the force acting on the actuation pin.

2.3.2 Locking stroke configuration

In this configuration the forces acting between the pin and the helical groove can be represented as in Figure 6 where N is the contact force acting between the pin and the groove, F the force acting on the pin from the component to which is fitted, R the component of the contact force that spins the shaft and α the wrap angle of the groove.

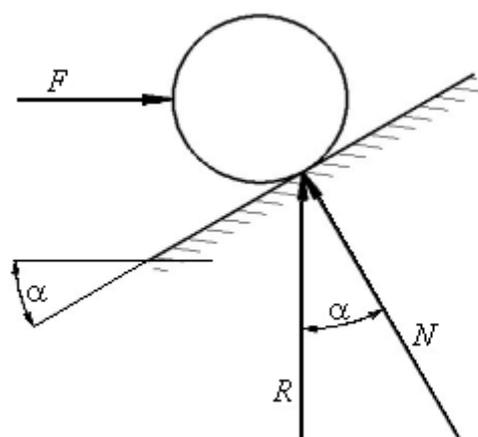


Figure 6 Forces acting on the pin.

The relations between the forces are:

$$\begin{aligned} N &= \frac{F}{\sin \alpha} \\ R &= N \cdot \cos \alpha = \frac{F}{\tan \alpha} \end{aligned} \quad (1)$$

The command spring pushes forward the cursor and the command pin that spins the shaft. The shaft rotation pushes forward the actuation pin and the sleeve.

3. MECHANISM DESIGN

3.1 STATIC DESIGN AND FATIGUE LIFE VERIFICATION

The design requires to size at first statically the mechanism: each component must resist to the maximum stress during exercise. This means that the maximum equivalent normal stress must be $\sigma_{id,max}$ lower than the yield strength of the material:

$$\sigma_{id,max} \leq R_Y \quad (2)$$

Once determined the relation between the equivalent normal stress and the component dimensions, these can be evaluated opportunely in order to satisfy relation (2).

Fatigue life verification requires a more complex approach and starts after the static design in which the component dimensions are set. There are several descriptions of the fatigue phenomenon, one of these states that the stress cycle can be represented by its mean stress and stress amplitude [11]. Haigh diagram represents the fatigue behaviour with respect to the combination of mean stress and alternate stress. Therefore once fixed the fatigue limit required, the Haigh diagram can be drafted and the analysed stress cycle can be positioned as a point in the graph: if the point is under the limit the component would satisfy the life requirement, if the point is above the limit the component would not. If the Haigh diagram is plotted with different life limits, once determined the stress cycle point it can be read the life of the component.

However a component can be stressed with different stress cycles along its life and each different stress cycle represents a different point on the Haigh diagram determining different fatigue lives [12].

By introducing the concept of damage, Miner law states how to determine the fatigue life of a component when stressed with different stress cycles. The damage is the ratio between the number of reiterations that a stress cycle acts on the component and the fatigue life of the component if it is stressed only with that specific stress cycle. The fatigue failure of the component occurs when the damage reaches the unitary value:

$$D = \sum_i \frac{n_i}{N_i} \quad (3)$$

where n_i is the number of cycles of the i -th stress cycle, and N_i the number of the i -th cycles that causes failure [11].

Stress cycles that are under the infinite life stress do not affect the component life or rather they do not contribute to the damage. Therefore cycles must be calculated with respect to the mechanism cycle. In this paper, cycles will be counted according to the peak-valley method [11].

Once fixed the materials, and then the Haigh diagram, the fatigue verification reduces to the determination of the mean stress and the alternate stress of each stress cycle. In facts, once determined the stress cycles coordinates on the Haigh diagram, the verification can be completed

evaluating the safety factor with respect to the fatigue limit fixed by requirements.

Figure 7 illustrates how the Haigh diagram can be built and how the verification can be done. The subscript U is referred to the ultimate tensile strength, Y to the yield strength, lim indicates the fatigue limit strength, a is referred to the stress amplitude and m to the stress mean. Point C is the stress cycle representation, while 1 and 2 are two fatigue limits according to different “working line” [12, 13].

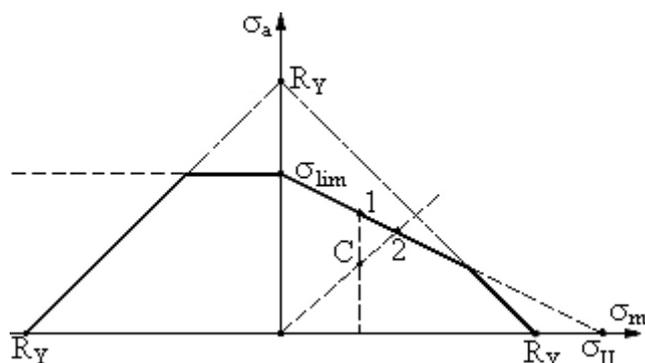


Figure 7 Haigh diagram.

There are mainly two different ways to evaluate the safety factor, depending on the influence to the failure of the only stress amplitude of the cycle or of both the mean stress and the amplitude. With respect to Figure 7, the fatigue safety factors are defined as:

$$S_{i,1} = \frac{\sigma_{a,1}}{\sigma_C} \quad (4)$$

$$S_{i,2} = \frac{\sqrt{\sigma_{a,2}^2 + \sigma_{m,2}^2}}{\sqrt{\sigma_{a,C}^2 + \sigma_{m,C}^2}}$$

During the Haigh diagram construction, it must take care of the modifying factors that take into account the difference between the component and the specimens (with fatigue limit σ_D) used for the Wohler diagram definition:

$$\sigma_{lim} = \sigma_D \cdot C_1 \cdot C_2 \cdot C_3 \cdot \square \quad (5)$$

There are various factors influencing the fatigue life such as: component dimensions, surface roughness, load type, surface treatment, temperature and the environment. The factors values can be found in literature as tables, graphs and formulas [12-14].

3.2 STRESSES DETERMINATION

The static design and the fatigue verification require very different determinations but both require the determination of the stresses under exercise. The static design considers only the maximum stress while the fatigue verification considers the mean stress and the stress amplitude. In the following paragraphs these stresses are determined. They are related to each other, as the maximum stress equals the sum of the mean stress and the stress amplitude:

$$\sigma_{max} = \sigma_m + \sigma_a \quad (6)$$

3.2.1 Command pin

The command pin is substantially stressed to shear by the normal force on the related groove. The maximum shear stress on a circular section during exercise is [13, 15]:

$$\tau_{max} = \frac{4}{3} \cdot \tau_{mean} \quad (7)$$

The maximum shear stress on the command pin is then:

$$\tau_{max} = \frac{16 \cdot N_l}{3 \cdot \pi \cdot d_p^2} = \frac{16 \cdot F_1}{3 \cdot \pi \cdot d_p^2 \cdot \sin \alpha_1} \quad (8)$$

The following relation, with null normal stress, can reduce the tangential stress to a normal equivalent stress, according to [16]:

$$\sigma_{id} = \sqrt{\sigma^2 + 3 \cdot \tau^2} = \sqrt{3} \cdot \tau_{max} \quad (9)$$

and then:

$$\sigma_{id} = \frac{16 \cdot F_1}{\sqrt{3} \cdot \pi \cdot d_p^2 \cdot \sin \alpha_1} \cong 2.94 \cdot \frac{F_1}{d_p^2 \cdot \sin \alpha_1} \quad (10)$$

Force F_1 is applied by a spring therefore it varies linearly between a range that depends on the stiffness of the command spring, on the total stroke of the BMDHS and on the initial deformation of the spring. This cycle is completed once in one mechanism cycle:

$$\sigma_m = 2.94 \cdot \frac{k_1}{d_p^2 \cdot \sin \alpha_1} \cdot \left(\Delta l_0 + \frac{c_T}{2} \right) \quad (11)$$

$$\sigma_a = 1.47 \cdot \frac{k_1 \cdot c_T}{d_p^2 \cdot \sin \alpha_1}$$

3.2.2 Actuation pin

During the rapid stroke, the actuation pin is not stressed while during the locking stroke, similarly to the command pin, it is stressed to shear by the elastic reaction of the actuation spring that varies linearly between the spring preload and the maximum force applied by the gas on the breechblock during combustion. According to the damage theory, the rest condition does not affect the fatigue life; therefore this condition can be neglected.

The stress cycle on the actuation pin is completed once in one mechanism cycle:

$$\sigma_m = 1.47 \cdot \frac{F_{gas} + k_1 \cdot \Delta l_0}{d_p^2 \cdot \sin \alpha_1}$$

$$\sigma_a = 1.47 \cdot \frac{F_{gas} - k_1 \cdot \Delta l_0}{d_p^2 \cdot \sin \alpha_1} \quad (12)$$

3.2.3 Sleeve

The sleeve is stressed by compression in the first half, between the actuation pin and the central inner annular rib

during the locked configuration, and by traction in the second half, during the rest configuration by the actuation spring preload.

The compression stress is:

$$\sigma_{N,C} = \frac{F_2}{\frac{\pi}{4} \cdot (d_e^2 - d_i^2)} \quad (13)$$

The compression mean stress and the amplitude are:

$$\sigma_{N,C,m} = 2 \frac{F_{\text{gas}} + k_2 \cdot \Delta l_0}{\pi \cdot (d_e^2 - d_i^2)} \quad (14)$$

$$\sigma_{N,C,a} = 2 \frac{F_{\text{gas}} - k_2 \cdot \Delta l_0}{\pi \cdot (d_e^2 - d_i^2)}$$

The tensile stress is:

$$\sigma_{N,T} = \frac{F_r}{\frac{\pi}{4} \cdot (d_e^2 - d_i^2)} = \frac{4 \cdot k_2 \cdot \Delta l_0}{\pi \cdot (d_e^2 - d_i^2)} \quad (15)$$

The mean stress and the stress amplitude are:

$$\sigma_{N,T,m} = \frac{2 \cdot k_2 \cdot \Delta l_0}{\pi \cdot (d_e^2 - d_i^2)} \quad (16)$$

$$\sigma_{N,T,a} = \frac{2 \cdot k_2 \cdot \Delta l_0}{\pi \cdot (d_e^2 - d_i^2)}$$

The inner rib is shear stressed by the action of the spring. The maximum shear stress acting on the rib is [13]:

$$\tau_{\max} = \frac{3}{2} \cdot \tau_m = \frac{3 \cdot F_2}{2 \cdot \pi \cdot d_1 \cdot s} \quad (17)$$

Therefore the equivalent normal stress is [16]:

$$\sigma_{\text{id}} = \sqrt{3} \cdot \frac{3 \cdot F_2}{2 \cdot \pi \cdot d_1 \cdot s} \cong 0.827 \cdot \frac{F_2}{d_1 \cdot s} \quad (18)$$

The mean stress and the stress amplitude on the rib root are:

$$\sigma_{r,m} = 0.413 \cdot \frac{F_{\text{gas}} + k_2 \cdot \Delta l_0}{d_i \cdot s} \quad (19)$$

$$\sigma_{r,a} = 0.413 \cdot \frac{F_{\text{gas}} - k_2 \cdot \Delta l_0}{d_i \cdot s}$$

All of these stress cycles occur once per mechanism cycle.

3.2.4 Breechblock

The breechblock (Figure 8) is stressed by shear at its rear rib that holds it inside the sleeve pushing on the locking ring nut.

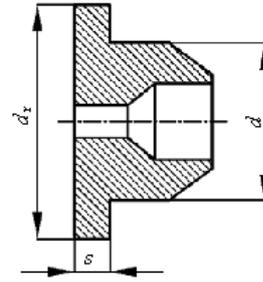


Figure 8 Sketch of the breechblock.

The shear stress is produced by the actuation spring preload when the mechanism is opened, decreases linearly during locking stroke, until it is not stressed when the breechblock rib is no longer in contact with the locking ring nut. The maximum shear stress is:

$$\tau_{\max} = \frac{3}{2} \cdot \tau_m = \frac{3}{2} \cdot \frac{k_2 \cdot \Delta l_0}{\pi \cdot d \cdot s} \quad (20)$$

The ideal stress is then [16]:

$$\sigma_{\text{id}} = \sqrt{3} \cdot \tau_{\max} \cong 0.827 \cdot \frac{k_2 \cdot \Delta l_0}{d \cdot s} \quad (21)$$

where d is the diameter of the breechblock and s the thickness of the breechblock rib.

The mean stress and the stress amplitude are then:

$$\sigma_m = \frac{0.413}{2} \cdot \frac{k_2 \cdot \Delta l_0}{d \cdot s} \cong 0.207 \cdot \frac{k_2 \cdot \Delta l_0}{d \cdot s} \quad (22)$$

$$\sigma_a = \frac{0.413}{2} \cdot \frac{k_2 \cdot \Delta l_0}{d \cdot s} \cong 0.207 \cdot \frac{k_2 \cdot \Delta l_0}{d \cdot s}$$

This stress cycle occurs one time each mechanism cycle.

3.1.5 Shaft

The shaft is stressed by the action of pins in two ways:

1. Torsion stress due to the tangential action on the grooves;
2. Pitting due to the contact stresses between the pins and the grooves [17].

The torsion on the shaft generates a tangential stress that is difficult to analyse because the grooves introduce a stress concentration near themselves and because the shaft has a complex cross section. The fatigue verification has to be performed on both the grooves, but the procedure is the same.

The torque acting on the shaft is constant on all its length and the value can be evaluated as:

$$T = R_1 \cdot \frac{d_1}{2} = \frac{F_1}{\tan \alpha_1} \cdot \frac{d_1}{2} \quad (23)$$

where T is the torque, d_1 the mean diameter of the helix of the groove. The subscript 1 indicates that the quantities are referred to the command pin side; a similar relation would have been if quantities were referred to the actuation pin side, with subscript 2.

The force generated by the command spring varies as described in §3.1.1; therefore the mean torque is:

$$T_m = \frac{k_1 \cdot \left(\Delta l_0 + \frac{c_T}{2} \right)}{\tan \alpha_1} \cdot \frac{d_1}{2} \quad (24)$$

and the torque amplitude is:

$$T_a = \frac{k_1 \cdot \frac{c_T}{2}}{\tan \alpha_1} \cdot \frac{d_1}{2} \quad (25)$$

To determine the stress state of the shaft there are literature tables of various sections, but it is difficult to evaluate the actual polar moment; thus it can be determined by the superposition of the outer shape polar moment with the inner cylindrical hollow (Figure 9).

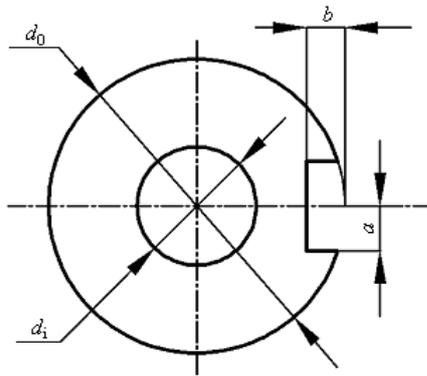


Figure 9 Cross section of the shaft.

The formula for the polar moment of the shaft is [18]:

$$J_s = J_o - J_i = C \cdot \frac{d_o^4}{8} - \frac{\pi \cdot d_i^4}{32} \quad (26)$$

where C can be determined as [18]:

$$C = K_1 + K_2 \cdot \frac{b}{d_o/2} + K_3 \cdot \left(\frac{b}{d_o/2} \right)^2 + K_4 \cdot \left(\frac{b}{d_o/2} \right)^3 \quad (27)$$

and the K_i factors are respectively [18]:

$$\begin{aligned} K_1 &= 0.7854 \\ K_2 &= -0.0848 + 0.1234 \cdot \frac{a}{b} - 0.0847 \cdot \left(\frac{a}{b} \right)^2 \\ K_3 &= -0.4318 - 2.2 \cdot \frac{a}{b} + 0.7633 \cdot \left(\frac{a}{b} \right)^2 \\ K_4 &= -0.0780 + 2.0618 \cdot \frac{a}{b} - 0.5234 \cdot \left(\frac{a}{b} \right)^2 \end{aligned} \quad (28)$$

The relations above are concrete if [18]:

$$\begin{aligned} 0.3 \leq \frac{a}{b} \leq 1.5 \\ 0 \leq \frac{b}{d_o/2} \leq 0.5 \end{aligned} \quad (29)$$

The maximum of the shear stress is reached at the outer surface of the shaft, and therefore [13]:

$$\tau_{\max} = \frac{T}{J_s} \cdot r_{\max} = \frac{T}{C \cdot \frac{d_o^4}{4} - \frac{\pi \cdot d_o^4}{16}} \cdot d_o \quad (30)$$

The stress concentration factor of this kind of keyway can be assumed equal to 1.7 [19].

The peak of the equivalent normal stress [16, 19] is:

$$\sigma_{id} = \sqrt{3} \cdot 1.7 \cdot \tau_{\max} \cong 2.94 \cdot \tau_{\max} \quad (31)$$

The mean and the alternate stresses can be evaluated inserting into (31) the tangential stress evaluated in (30) with the mean (24) and amplitude (25) torque:

$$\begin{aligned} \sigma_m &= 2.94 \cdot \frac{d_o}{C \cdot \frac{d_o^4}{4} - \frac{\pi \cdot d_i^4}{16}} \cdot \frac{k_1 \cdot \left(\Delta l_0 + \frac{c_T}{2} \right)}{\tan \alpha_1} \cdot \frac{d_1}{2} \\ \sigma_a &= 2.94 \cdot \frac{d_o}{C \cdot \frac{d_o^4}{4} - \frac{\pi \cdot d_i^4}{16}} \cdot \frac{k_1 \cdot c_T}{\tan \alpha_1} \cdot \frac{d_1}{4} \end{aligned} \quad (32)$$

This cycle runs once per mechanism operational cycle.

The hertzian stress generated by the contact between the cylindrical pins and the grooves is [13, 15]:

$$\sigma_H = 0.564 \sqrt{\frac{N}{h_G \cdot \Delta} \left(\frac{1}{R_1} + \frac{1}{R_2} \right)} \quad (33)$$

where N is the contact force, h_G the depth of the grooves, R_1 the pin radius and R_2 the groove radius which tends to infinite; Δ is a function of the elasticity modulus E and the Poisson ratio ν , of the two coupled materials [16]:

$$\Delta = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \quad (34)$$

Therefore, if the materials are the same, they have the same Poisson ratio and Young modulus, and then:

$$\sigma_H = 0.564 \sqrt{\frac{N}{2 \cdot h_G \cdot R_1} \cdot \frac{E}{1 - \nu^2}} \quad (35)$$

If the hertzian stress is lower than the maximum contact stress admissible for the used material, the component will not fail before the required life.

4. CONCLUSIONS

The present paper illustrates the design procedure of the principal components of the breech mechanism with double helix shaft.

At first each component must be designed statically, calculating opportunely its critical dimensions. Then it must be verified to the fatigue phenomenon. This verification procedure requires to build the Haigh diagram, even if not in a graphical way, but also in an analytic way, or even building a program or a subroutine that, from the input coordinates of mean stress and stress amplitude, can evaluate the safety factor.

Each stress cycle then must be analysed at first determining the kind of stress, its relation with the forces acting inside the mechanism, and then inserting the force variation during the mechanism work. Once determined the stress cycle, it must be determined the mean stress and the stress amplitude in order to position the stress cycle on the Haigh diagram to determine the safety factor.

The verification procedure is completed reiterating for each stress cycle of each component.

The evaluated stress cycles occur each one time per mechanism cycle, therefore the mechanism fatigue life is strictly correlated to the first fatigue failure of the components.

REFERENCES

- [1] Quinian J.B. and Vetter E., *Caseless ammunition cartridge*, U.S. Patent 3,557,700, 1969.
- [2] Flanagan J.E. and Lo G.A., *Caseless ammunition*, U.S. Patent 4,187,781, 1972.
- [3] Goldin M., *Fully telescoped caseless cartridge*, U.S. Patent 3,557,700, 1970.
- [4] Gorka S., Outer Limited, *American Handgunner*, No. 5, 2001.
- [5] Kessner B.C., Marine Corps Still Eyeing Caseless Rifle Possibilities, *Defense Daily*, No. 10, 2009.
- [6] Delprete C., Bonisoli E. and Tomasi D., Armi “ad hoc” con la progettazione parametrica, *Il progettista industriale*, No. 6, 2012.
- [7] Price R.M. and Reed R.Jr., *Thermally stable gun and caseless cartridge propellants*, U.S. Patent 4,263,070, 1974.
- [8] Braak E.C. and McCarty K.P., *Surface coated smokeless propellant charge for caseless ammunition having heat resistant properties*, U.S. Patent 3,713,916, 1972.
- [9] Tomasi D., *Meccanismo otturatore con albero a doppia elica*, I Patent TO2011A001061, 2011.
- [10] Tomasi D., Parametric design of an automatic gun system with caseless ammunition, *Master Degree Thesis*, Politecnico di Torino, 2012.
- [11] Lee Y.L., Pan J., Hathaway R. and Barkey M., *Fatigue testing and analysis - Theory and practice*, Elsevier, 2005.
- [12] Bannantine J.A., Comer J.J. and Handrock L.J., *Fundamentals of metal fatigue analysis*, Prentice Hall, 1989.
- [13] Juvinall R.C. and Marshek K.M., *Fundamentals of machine component design*, John Wiley & Sons, 2006.
- [14] Campbell F.C., *Elements of Metallurgy and Engineering Alloys*, ASM International, 2008.
- [15] Budynas R.G. and Nisbett J.K., *Shigley's Mechanical Engineering Design*, McGraw Hill, 2006.
- [16] Liu J. and Zenner H., Fatigue limit of ductile metals under multiaxial loading, *Biaxial/Multiaxial Fatigue and Fracture*, Vol. 31, No. 8, pp. 147-164, 2003.
- [17] Glaeser W.A. and Shaffer S.J., Contact Fatigue, *ASM Handbook*, Vol. 19, No. 27, pp. 331-336, 1996.
- [18] Young W.C. and Budynass R.G., *Roark's Formulas for Stress and Strain*, McGraw-Hill, 2002.
- [19] Seely F.B. and Dolan T.J., Stress concentration at fillets, holes, and keyways as found by the plaster-model method, *University of Illinois Bulletin*, Vol. 32, No. 42, 1935.

MODULAR RECONFIGURABLE ANTHROPOMORPHIC GRIPPER WITH THREE FINGERS FOR ROBOTS-DESIGN, PROTOTYPE AND SIMULATION

Catalin Constantin Moldovan Ionel Staretu

Transilvania University of Brasov, Romania

ABSTRACT

One anthropomorphic modular reconfigurable gripper, with three fingers, for robots is described in this paper. For the first time the stages of synthesis, analysis design and functional simulation are presented. The structural synthesis of the anthropomorphic grippers for robots can be made regarding the following main criteria: the number of fingers, the number of phalanxes, the relative dimensions of the phalanxes, the relative position of the fingers, the degree of freedom of the gripping mechanism and the characteristic constructive elements used. We choose a version with three identical fingers with three phalanxes on each finger. Correct closing of the finger and of the gripping mechanism is obtained from kinematic synthesis. The static synthesis solves the problem of obtaining the necessary gripping force on each finger and the total gripping force. By means of constructive dimensions a 3D model can be obtained using CATIA soft. For one variant of this type of gripper, the technical documentation is completed and the technical project has all the conditions for practical achievement and a prototype was made. The main constructive modules are: the support – the palm and the finger. Main technical characteristics of the prototype are indicated.

Keywords: Grasp, Anthropomorphic gripper, Reconfigurable

1 INTRODUCTION

The anthropomorphic grippers are of increased interest due to raising applicability to industrial robots but also to other types of robots, especially humanoids service robots. Currently there is a relatively large variety of such grippers ([6]; [2]; [3]; [18]; [7]; [4]; [1]), which have a high price, even in some cases very high, which discourages attempts to introduce them in current applications.

The paper first briefly refers to the class of modular reconfigurable anthropomorphic grippers proposed by the authors.

In terms of modularization, this class is based on a generic version, which may have a variable number of fingers. This number ranges from 2 to 6, but there is as well the opportunity to make a gripper in a wide range of sizes, from small sizes (0.75, 0.5, 0.25, etc. reported to the human hand) to larger versions (1.5, 2, 2.5, etc. reported to the human hand).

The possibility of being reconfigurable refers to the use of a platform where fingers (three or more) can have more relative positions only through disassembling and assembling elsewhere, without removing the platform off the robot arm.

In this paper for a variant of this class of anthropomorphic grippers, that is a gripper with three fingers, the main theoretical and construction features are illustrated and a prototype is described as well. Some aspects regarding CAD simulation and virtual simulation with two methods are shown too.

Obviously, all considerations can be extrapolated to variants with fewer fingers, respectively, two or more fingers, four, five or even six.

2 STRUCTURAL SYNTHESIS AND ANALYSIS

2.1 STRUCTURAL SYNTHESIS

The structural synthesis seeks to set possible configurations and the structure of a finger so that it has the largest degree of utility possible.

Contact author: Catalin Constantin Moldovan

E-mail: catalin.moldovan@unitbv.ro

Looking at possible configurations four are considered significant (see Fig. 1), which can be obtained by proper installation of the three fingers on the same platform.

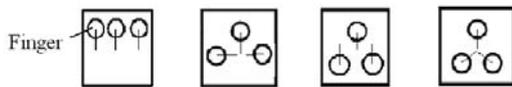


Figure 1 Significant configuration of three fingers.

In connection with the structure of a finger it may have two or three phalanxes (see Fig. 2), possibilities of which we opted for three phalanxes, for a greater degree of utility (see Fig. 2,b).

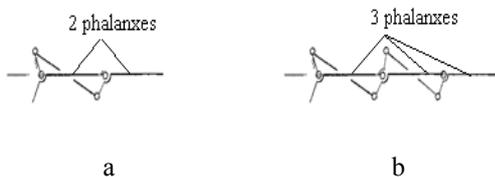


Figure 2 Fingers with two phalanxes (a) and three phalanxes (b).

There is the possibility of using four phalanxes too, or even five, which must be duly justified, however, as there are clearly higher prices.

2.2 STRUCTURAL ANALYSIS

The mechanism of the finger (see Fig. 3) is a poly-contour mechanism with two outside connection $L=2$ ($v_1, F_1; v_{P1}, FP_1$ – see Fig. 4,a) and degree of freedom $M=1$.

The degree of freedom is obtained with $M = \sum M_i - \sum f_c$, where M_i is the degree of freedom for mono-contour i mechanism and $\sum f_c$ is the degree of freedom for common joints (see Fig. 4,b).

For each mono-contour mechanism the degree of freedom is obtained with $M = \sum f_i - \chi_K$ (where $\sum f_i$ is the degree of freedom of the joints and χ_K is cinematic degree of the mono-contour k mechanism ([12])).

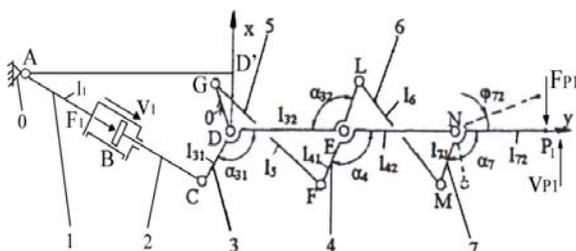


Figure 3 The structural scheme of the finger.

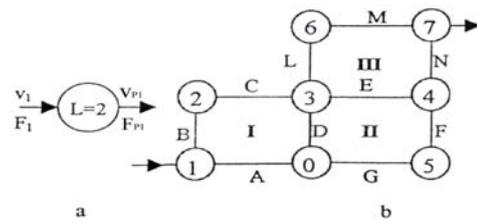


Figure 4 The block scheme and the graph of the mechanism.

For the mechanism shown in Fig. 3, in accordingly with the graph of Fig. 4,b, the following relations are obtained:

$$M_I = f_A + f_B + f_C + f_D - \chi_I = 1 + 1 + 1 + 1 - 3 = 1$$

$$M_{II} = f_D + f_E + f_F + f_G - \chi_{II} = 1 + 1 + 1 + 1 - 3 = 1 \quad (1)$$

$$M_{III} = f_L + f_M + f_N + f_E - \chi_{III} = 1 + 1 + 1 + 1 - 3 = 1,$$

$$\text{and } \sum f_c = f_D + f_E = 1 + 1 = 2.$$

The degree of freedom will be:

$$M = M_I + M_{II} + M_{III} - \sum f_c = 1 + 1 + 1 - 2 = 1 \quad (2)$$

$M=1$ has the following significance: one independent movement (speed): $v_1 = s_1$ and one function of the external forces: $F_1 = F_1(F_{P1})$. $L-M = 1$ represents one function of movement: $v_{P1} = v_{P1}(v_1)$ and one independent force: FP_1 – the contact force between finger and grasped object.

3 KINETOSTATIC SYNTHESIS AND ANALYSIS

3.1 KINETOSTATIC SYNTHESIS

In this phase we determine linear and angular dimensions of components so that the fingers close properly (kinematic synthesis purpose), and the given weight can be gripped and handled (static synthesis purpose).

3.2 KINEMATIC ANALYSIS

The function of position, the function of speed and the function of acceleration for characteristic P_i points are obtained from the cinematic analysis. The vectorial close chain method is used successively for each mono-contour mechanism. The vectorial equations are:

$$\overline{AC} + \overline{CD} + \overline{DD'} + \overline{D'A} = 0 \quad (\text{see Fig. 5a}),$$

$$\overline{DE} + \overline{EF} + \overline{FG} + \overline{GD} = 0 \quad (\text{see Fig. 5b}) \text{ and}$$

$$\overline{EN} + \overline{NM} + \overline{ML} + \overline{LE} = 0 \quad (\text{see Fig. 5c}) \quad ([12], [13]).$$

The implicit form for the equation of position is: $\varphi_{72i} = \varphi_{72i}(s_i)$, i – the number of the fingers: $i=1,2,3$.

The functions for speeds are the derivative function of time of the functions for positions and the functions for accelerations are the derivative of the functions for speeds:

$$v_{P_i} = \dot{\varphi}_{72i}, \quad a_{P_i} = \ddot{\varphi}_{72i} \quad (3)$$

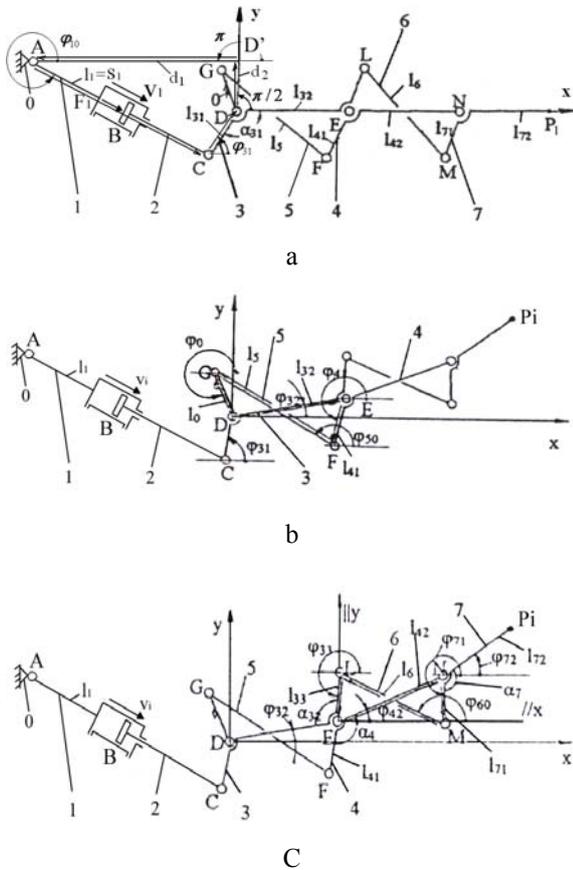


Figure 5 The cinematic schemes.

3.3 STATIC ANALYSIS

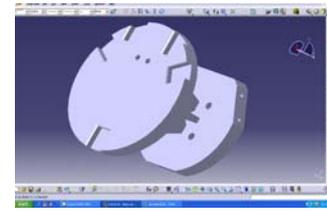
The function of the external forces is obtained from the theorem of balance between the powers of entrance and emergence of mechanism, $v_i \cdot F_i + v_{Pi} \cdot F_{Pi} = 0$, and

$$F_i = - \frac{v_{Pi} \cdot F_{Pi}}{v_i} \quad (4)$$

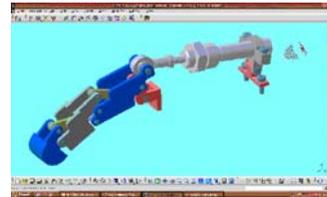
The internal forces are calculated using the theorem of the joints and, afterwards, with the balance static equations of the mobile elements ([11], [13]).

4 CONSTRUCTIVE DESIGN, 3D MODEL AND PROTOTYPE

The calculation of strength was made in function of the internal forces which act between elements. With the constructive dimensions a 3D model can be obtained using CATIA soft. There are two main constructive modules: the support – the palm (see Fig. 6,a) and the finger, see Fig. 6,b, ([11], [12], [13], [14], [15]).



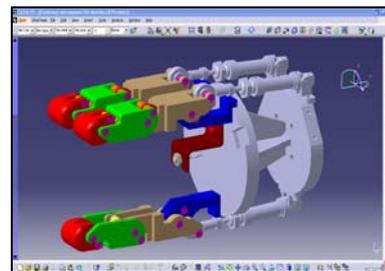
a



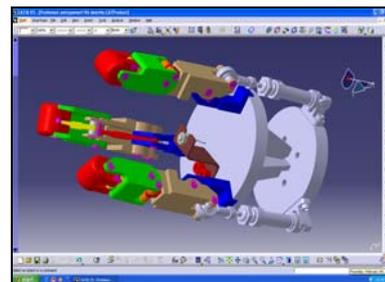
b

Figure 6 The main constructive modules.

With these modules two main three-finger versions can be obtained (see Fig. 1), the fingers having possible parallel (see Fig. 7a) or concurrent movements (see Fig. 7b).



a



b

Figure 7 Modular anthropomorphic grippers with three fingers.

On the basis of the technical documentation prepared in accordance with technical rules in force a prototype of the gripper analyzed in this paper was issued (see Fig. 8,a). In Fig. 8,b, as a first experimental form and functional testing, gripping a spherical body with this prototype is exemplified.

The main technical characteristics of this prototype are: degree of freedom: $M=3$; weight hand: 12 N; payload: 40 N; gripping force: ~ 30 N/finger; dimensions: finger : 1:1 human fingers size and hand: 140x140x100 mm.



a



b

Figure 8 Modular anthropomorphic reconfigurable three-fingered gripper prototype.

This gripper, with one specific intermediary piece, can be mounted on any industrial commercial robot (see CAD simulation in Fig. 10). One of its configurations can be obtained, during the gripper is mounted on robot, with change the relative position of the fingers only, regarding the form of the grasped object.

For the drive, pneumatic linear motors are used. The prototype can be equipped with contact sensors (for example, type CZN-CP15), and command and control can be ensured through appropriate equipment. After providing corresponding equipment, the prototype will be mounted on a robot and will be fully tested in various gripping operations, including handling.

5 CAD SIMULATION AND VIRTUAL SIMULATION

5.1 CAD SIMULATION

A functional CAD simulation (see Fig. 9) was made to check the correct work and to identify the solutions to obtain the optimum variant for this gripper. A CAD simulation with the gripper mounted on robot is shown in Fig. 10.

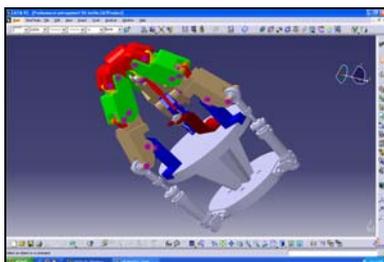


Figure 9 CAD functional simulation of the gripper.



Figure 10 CAD example with the gripper mounted on robot

5.2 VIRTUAL SIMULATION

The general structure of the virtual system ([8]) can be adapted as well to cases when there is an anthropomorphic gripper for robots in the virtual environment, based on five or less fingers. This structure particularization when a three-finger gripper is placed in the virtual scene, respectively, our anthropomorphic gripper is illustrated in Fig. 11 ([9]).

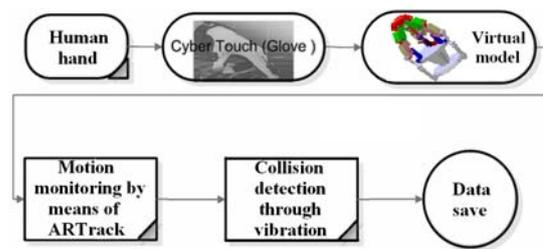


Figure 11 Virtual grip system structure.

A development of the application is to integrate a communication module based on Cyber Touch to read rotations at the level of the user's fingertips.

The values thus obtained correspond to direct kinematics of the virtual gripper.

Virtual simulation achievement means to detect the collision between the virtual gripper and the entities in the virtual scene, which can be made using an OPCODE (Optimized Collision Detection) library, thus gripping virtual simulation structure configuration was achieved (see Fig. 12).

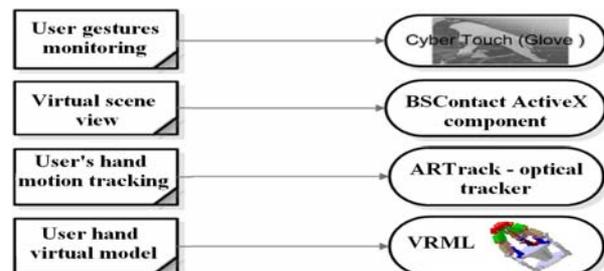


Figure12 Implementing particularities.

The main result of this simulator is gripping in virtual environment of objects and data transfer for real gripper fingers pre-configuration to grip a real object identical to the virtual one.

Gripping optimizing means adding multimodal interfaces to communicate with the gripper. A scheme for such interface is described in ([9]), which facilitates Barrett gripper closing after pre-configuration, after the part transfer, open by verbal command, which means direct interaction of three components: speech recognition, remote handling and direct interaction with the gripper, possible with the use of a PDA/Smartphone mobile device ([9]). This scheme can be easily adapted for our reconfigurable gripper.

Another method, now in progress, proposes an architecture for a system that detects human hand gestures in real time, and uses these gestures to control the virtual gripper ([5]). An overview of the proposed architecture for this system is shown in Fig 13.

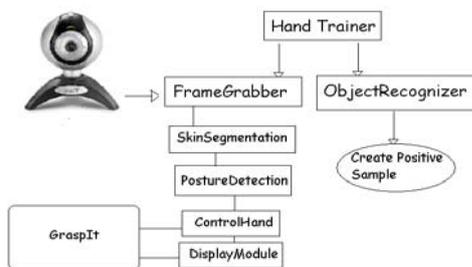


Figure 13 Application system design.

As it can be seen, the system consists of five modules, namely FrameGrabber, SkinSegmentation, PostureDetection, ControlHand, DisplayModule, summarized in the next paragraph.

In the “Vision Based” hand gestures detection systems, hand movements are recorded by a video camera. The algorithms of the system are executed separately on each frame from the entire video sequence. To gain speed, before executing an algorithm, an initial filter that eliminates unnecessary data and highlights necessary data is applied.

Initialization

To implement the hand detection and gesture recognition module an image processing library called OpenCV was used.

OpenCV Library is free and can run on multiple operating systems (Windows / Linux / MacOS).

Using this library, *HandTrainer* application was created. This application uses the digital web camera to capture images, and creates a set of positive and negative sample images which are used to train Haar classifiers. The training process needs to be executed only once, after this, Haar classifiers can be used anytime to detect human hand from a video stream.

Implementation

The *Frame Grabber* module is used to capture all images from a stream. This can represent a part from a video file, or directly the image stream from the webcam. On this module, each frame, which later will be processed and contain human hand gesture, represent the output.

The *SkinSegmentation* module uses as input data, the result image from the FrameGrabber module. On each image, the SkinSegmentation module will detect the regions that contain skin's color. The method used to detect skin images is presented in ([18]) at a conceptual level.

The output from this module represents the same image from FrameGrabber together with one array containing all pixels region where skin was detected.

PostureDetection module uses the array from SkinSegmentation in which Haar classifiers trained in the CreateSample module are used to detect the hand posture. Currently the system has been implemented to detect gestures for the fist and open hand (see Fig. 14).



Figure 14 Fist and open hand.

The *ControlHand* module will generate input data for the control of a virtual hand in GraspIT simulator.

DisplayModule uses the GraspIT application presented in ([10]) to simulate virtual hand (see Fig 15,a,b) and similar virtual gripper (see Fig. 15,c).

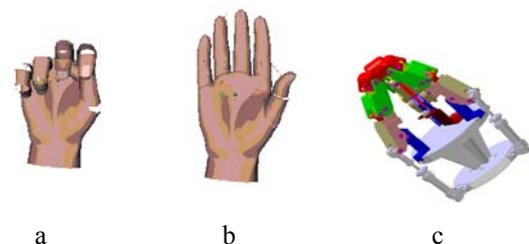


Figure 15 Simulated virtual hand (fist and open hand: a, b) and simulated virtual gripper (fist gripper: c).

For functional simulation of the grasped operations, the robot with the gripper were transferred in virtual reality – VRML soft and its become a virtual robot and a virtual gripper (see Fig 16). Here we can test different grasping operations for different objects. Now we test the second method above mentioned. Then, the results, for one correct grasp, can be used for programming the real gripper.



Figure 16 The robot with gripper in VRML soft.

If in the first stage testing will be done in CAD environment (including gripping phase), test done in a preliminary stage without the object to grip (Fig. 9 and Fig.10), then functional simulation will be possible in virtual environment (e.g. VRML) in order to establish data for accurate virtual gripping and their transmission to the real gripper – the prototype ([15], [16], [17]).

6 CONCLUSIONS

The next conclusions can be formulated in according to the aspects presented:

- a. The main stages for to design the anthropomorphic mechanical grippers are: structural synthesis and analysis, cinematic synthesis and analysis, static synthesis and analysis, constructive design and 3D model and functional simulation (CAD and virtual).
- b. The variants of the family of the mechanical anthropomorphic modular reconfigurable grippers for robots, with two, three, four and five identically fingers, can be obtained in accordance with the number and the relative position of the fingers.
- c. The modular reconfigurable grippers with three fingers can be obtained using two main modules: the support – the palm and the finger.
- d. CAD simulation and virtual simulation are very important for to obtain an optimum gripper and gripping operations.

ACKNOWLEDGEMENTS

The authors would like to acknowledge IBM Romania support and assistance for developing this research.

REFERENCES

- [1] Barrett. 2010. www . barrett.com/robot.
- [2] Butterfass J. et al., DLR-Hand II: Next Generation of a Dextrous Robot Hand. In: *Proc. of the IEEE Int. Conference on Robotics and Automation*, Seoul, Korea, 2001.
- [3] Butterfass J. et al., Design and Experiences with DLR Hand II. In: *Proceedings of the World Automation Congress*, Seville, Spain, 2004.
- [4] Ceccarelli M. et al., Design and tests of a three finger hand with 1-DOF articulated. *Robotica Journal*, Vol. 24, Issue 2, pp. 183-196, 2006.
- [5] Chen Q. *Real-Time Vision-Based Hand Tracking and Gesture Recognition*, Doctorate thesis, 2008.
- [6] Kawasaki H. and T. Komatsu. Mechanism Design of Anthropomorphic Robot Hand: Gifu Hand I. *Journal of Robotics and Mechatronics*, Vol.11, No. 4, pp. 269-273, 1999.
- [7] Kochan A. Shadow delivers first hand, *Industrial Robot Journal*, V. 32, Issue 1, pp. 15-16, 2005.
- [8] Itu A. *Contributions to Gripping Strategies in Real and Virtual Environment using a Three-Fingered Anthropomorphic Gripper*, Doctorate thesis, Transilvania University of Brasov, Romania, 2010.
- [9] Itu A. and I. Staretu. Multimodal Control Interface for an Anthropomorphic Gripper. In: *Proc. of 19th International Workshop in Alpe-Adria-Danube Region- RAAD 2010*, Budapest, Hungary, pp. 61-64, June 23-25, 2010.
- [10] Miller, A. Grasp it a versatile simulator, Doctorate thesis, U.S.A, 2006.
- [11] Starețu I., *Sisteme de prehensiune (Gripping Systems)*, Lux Libris P.H., Brașov, Romania, 1996.
- [12] Starețu, I. and I. Daj, *Mechanisms and Machine Elements (in Romanian)*, Lux Libris P.H., Brasov, Romania, 2000.
- [13] Starețu I. et al. Mechanical Hand. Anthropomorphic Gripping Mechanism for prostheses and robots (in Romanian), Lux Libris P.H., Brasov, Romania, 2001.
- [14] Staretu I., M. Ionescu and V. Runcan. Family of Mechanical Anthropomorphic Poly-Mobile Grippers for Robots – Synthesis, Analysis, Design and Functional Simulation. In: *Proc. of The 15th International Workshop on Robotics in Alpe-Adria-Danube Region-RAAD*, Balatonfured, Hungary, pp. 273-277, June 15-17 2006.
- [15] Starețu I., Familie de prehensoare antropomorfe pentru roboți (Family of Anthropomorphic Grippers for Robots), *Buletinul AGIR*, No.4, Year XII, pp. 36-42, 2007.
- [16] Staretu I. and Bolboe M., Synthesis, Analysis, Design and Functional Simulation for a Family of Anthropomorphic Grippers for Robots. In: *Proc. of the SISOM 2007*, Bucharest, Romania, pp. 131-134, 29-31 May 2007.
- [17] Starețu I., Anthropomorphic Gripping Systems with Jointed Bars or Wheel and wires for Industrial Robots – Constructive Synthesis, Analysis and Design, *New Trends in Mechanisms*, Editors Crețu S.M. and Dumitru N, Academica-Greifswald Publishing House, Germany, pp. 133-144, 2008.
- [18] Yogarajah P. et al. , A Dynamic threshold approach for Skin Segmentation in Color Images. In: *Proc. of 2010 IEEE 17th International Conference on Image Processing*, Hong Kong, September 26-29, 2010.

VELOMOBILES: DESIGN GUIDELINES

Paolo Baldissera

Cristiana Delprete

Matteo Tirelli

Department of Mechanical and Aerospace Engineering - Politecnico di Torino
Corso Duca degli Abruzzi 24 – 10129 Torino, Italy

ABSTRACT

The article provides the general guidelines to design a velomobile as a solution for the always-rising mobility problems. After a brief introduction the velomobile is technically defined, with a close look to the dimensions of the available models. A flowchart is provided with the minimum needed steps to design it and each step is quickly described. Particular attention is given to the lower limb movement size to design the fairing, by proposing a general formula to define it as first attempt. Breaking systems are analyzed, with the needed equations to choose it, as well as fairing ducts to ensure the needed airflow for the cyclist comfort. Dumpers, steering, transmission, pedaling assistance and fairing design are described, with the basic problems and possibilities for each part.

Keywords: velomobile design, urban mobility, sustainability.

1 INTRODUCTION

Human transport systems evolved quite a lot in the last century under the influence of two factors: technology and society. These two factors did not always have the same importance in history. Most of the times new technologies brought changes in societies where they were introduced. Over the years, societies stabilized and evolved around those innovations, becoming essential for their survival.

Nowadays an always-bigger sensitization to problems consequent to the use of modern conventional transport systems is rising: air and acoustic pollution, traffic congestion in cities, oil dependence and accidents frequency. Despite this awareness, society cannot easily change habits: a strong structural program promoted by authorities and a true, popular alternative is needed.

In this context velomobiles are quickly gaining ground as evolution of the popular bicycles. This transport system is increasingly appreciated through North Europe societies, while the rest of the world seems to ignore it, except from small groups of enthusiasts.

The true potential of this innovation to solve the above listed mobility problems is a fact, and this article would give the general guidelines to design a velomobile, investigating the different technical choices available to the manufacturer.

2 VELOMOBILE DEFINITION

A velomobile is a fully-enclosed recumbent trike. The driver position is the same as classic recumbent bikes, more comfortable than popular bicycles. The three wheels solution allows the velomobile to be driven easier than classic recumbent bicycles, solving the common balance problem in those types of vehicles and increasing agility and driveability. For a trike there are two possible types of configurations, depending on where the coupled wheels are located: tadpole or delta. Tadpole has two front wheels, while delta has two rear wheels. The structure itself of a velomobile makes it heavier than a classic bicycle and the three wheels increase the rolling resistance. These drawbacks are almost completely surpassed by the quite better aerodynamics offered by velomobiles in comparison with classic bicycles, thanks to the fairing that also improves driver protection from accidents and weather.

The three wheels solution may cause stability problems in quick steering, which are less important in the tadpole configuration and/or can be completely solved through tilting-steering systems. This last solution is not widespread because it is not easy to implement and it is also expensive. For urban purpose the non-tilting tadpole configuration quickly became the standard and the present article refers to this solution. An overview of the velomobile design approach is proposed in the following, with a special focus on some of the most common technical questions that can arise during the first layout definition.

Contact author: Matteo Tirelli

E-mail: matteo.tirelli@polito.it

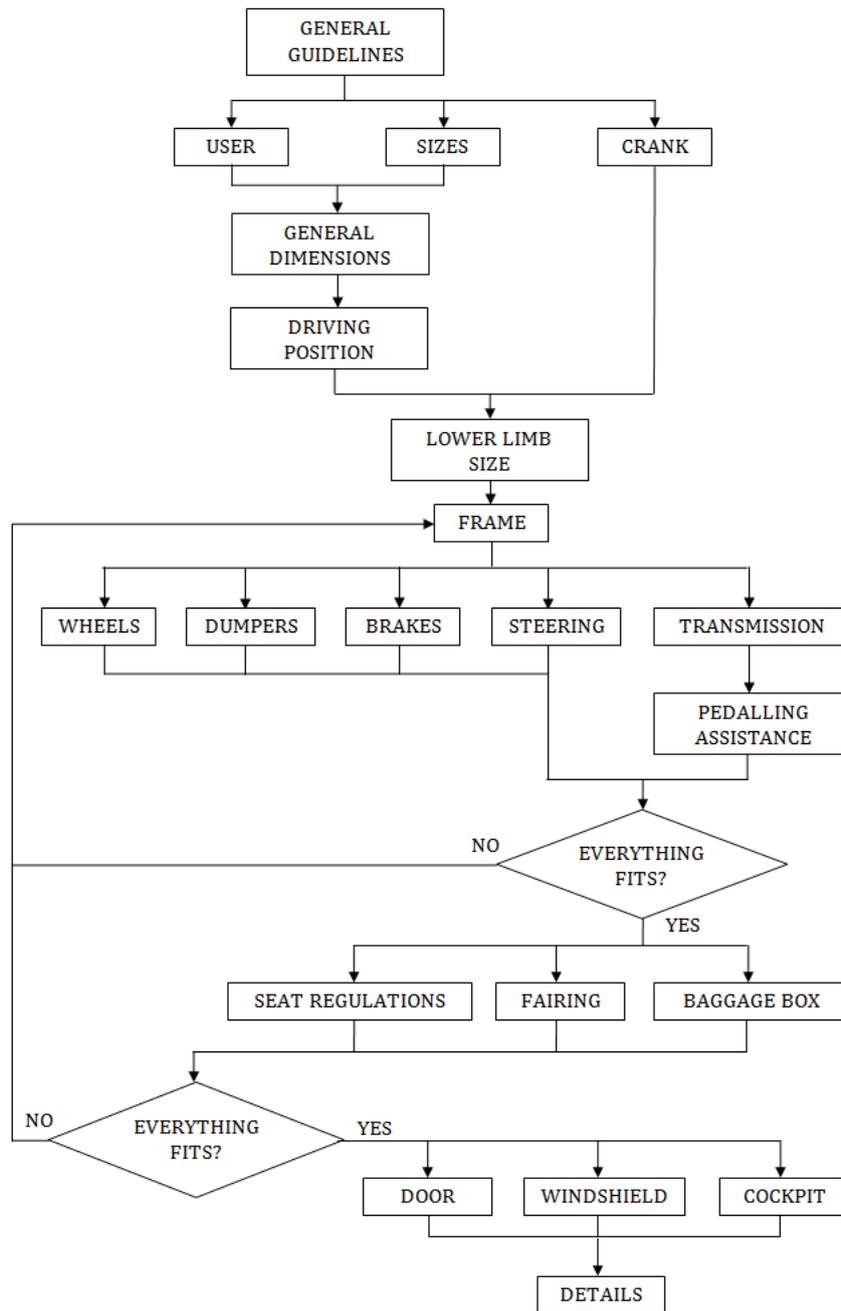


Figure 1 Flowchart proposal for velomobile design.

3 VELOMOBILE DESIGN

3.1 GENERAL CONSIDERATION

Almost forty velomobile models are available on market up to now, and they are very similar each other. The general shape is directly derived from racing prototypes built to take part at competitions. The daily use sacrifices maximum performances in order to improve driver comfort, introducing useful features like a baggage van as example.

The biggest limits of competition vehicles are the external visibility, due to the small height from the ground, and the accessibility; both can be solved by lifting up the velomobile centre of mass and by designing a proper access door.

The general steps to design a velomobile are represented by the flowchart in Figure 1. It is not a universal way to follow, but it summarizes and orders all the considerations one has to think about to properly design a velomobile.

The first step to design an urban velomobile is to define its user. From anthropometric data a human puppet can be defined, representing the physical dimensions of a reasonable percentage of the population, identified as the market target. Those dimensions would not include the whole population because it would need a wide range of regulations to think about: it is more convenient to build different sizes for each needs. Each size is adequate for a certain range of users and is designed around the dimension of the biggest suitable man: smaller ones from the same size use dedicated adjustments, introduced later in this article.

The next step is to choose the mountable crank length, not different from the standard dimensions commonly used in classic bicycles, and the general velomobile dimensions, strongly related to the project objectives. This only set the guidelines for the following design steps and the dimension may be modified as needed. Comparing the models available in the market, with public information, the typical dimensions are:

- Length: from 2000 mm to almost 2800 mm, average around 2600 mm (Figure 2, top);
- Width: from 600 mm to almost 1000 mm, average around 800 mm (Figure 2, middle);
- Height: it is really variable, from 800 mm up to 1300 mm, average around 1000 mm (Figure 3, bottom);
- Height from the ground: it cannot be too small for urban purpose, variable from 50 mm to almost 180 mm;
- Distance between front and rear wheels: it varies from 1200 mm to 1350 mm. Generally velomobiles are divided in long, medium and short wheel base: increasing wheelbase distance improves stability at higher speeds, while decreasing improves drivability, reducing the steering radius. The difference between those solutions is evident by looking the relative position of the bottom bracket from the front wheels, as shown in Figure 3. For the urban target a short wheelbase is preferable, as shown by the last market trend in recumbent sales [1];
- Distance between coupled wheels: it is not all that different from width, variable from 600 mm to 800 mm. It mainly affects the vehicle steering stability: increasing this distance improves the ability to run through quick corners, but also the velomobile size, an important factor for urban agility.

These general design guidelines, with the chosen driving position, define the maximum driver size from where to start modelling the cockpit. About the driving position, studies showing muscular activity [2] and cycling performances changing the driver seat angle [3] and the relative angle between bottom bracket and cyclist trunk [4], (known as BCA, Body Configuration Angle) are available in literature.

The suggested optimal trunk angle for a good performance is 75 degrees inclined from the vertical line passing through the bottom bracket [3], and it is also a good value for visibility, aerodynamics and comfort, while the optimal BCA angle is from 130 degrees to 140 degrees [4].

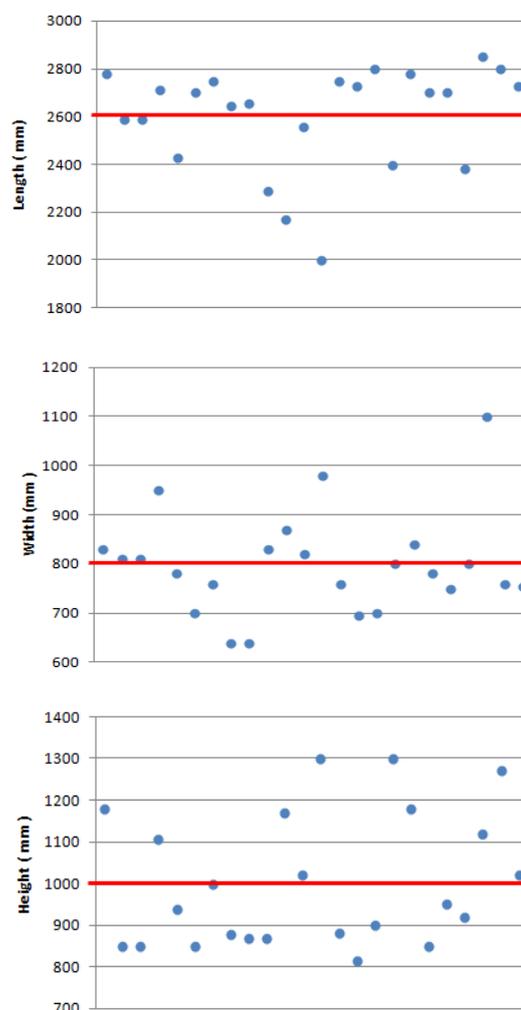


Figure 2 Market velomobile length, width and height (average values in red).

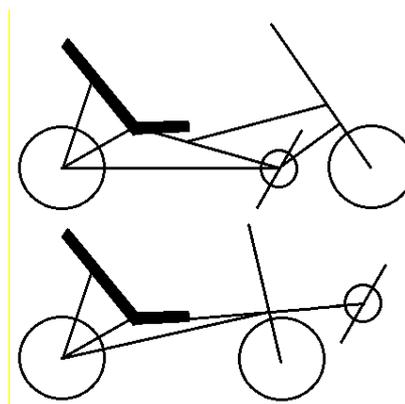


Figure 3 Long and short wheelbase layouts.

It is important to define the movement of the lower limbs during pedalling in order to know the maximum size around which to design the fairing, introducing also some safety margins to be sure the cyclist does not bump it, as shown in Figure 4.

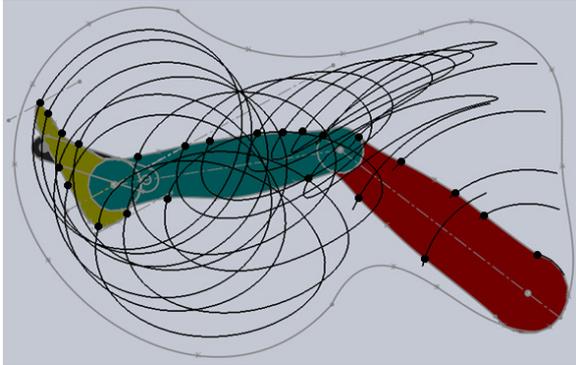


Figure 4 Movements of the lower limb during pedalling and volume occupied.

The margins are drawn by an offset of the more external curves resulting from the motion study. The offset value has to be a good compromise between safety and exaggeration: 20% of the chosen foot length is a good start value. The volume size width depends on the cyclist dimensions. Fixed dimensions are the distance between the two feet and the width of the thighs. These dimensions do not change so much from one to another, and for this reason a start width around 400 mm is normally enough. The shape of the size volume is not easy to reproduce. Keeping it to more extensive studies, it is possible to approximate the size with a shape easier to handle. While the other parts are properly replaced with straight lines, for the fairing aerodynamics it is important to keep the curved line around the feet movement. For this reason a circumference centred in the bottom bracket is drawn with radius R , connected with a rectangle $X \times Y$ containing the lower limb of the cyclist (Figure 5). The front downside line changes from the original shape, but this is not a problem because the wheel dimension, main constraint for the fairing design, also affected this size volume.

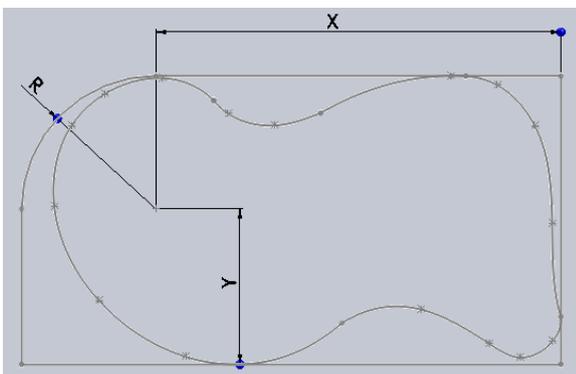


Figure 5 Simplified size volume.

Since the shape of the size volume depends on the percentage of the target cyclist chosen in the first design phase, some general relationships between this percentage and the reference dimensions may be obtained. Studying the size from 1%, 50% and 99% of the population dimensions is possible to approximate the design shape as:

$$R = 0.007289 p^2 - 0.7187 p + 283.7$$

$$X = -0.002082 p^2 + 2.412 p + 686.6 \quad (1)$$

$$Y = 0.001041 p^2 + 0.3551 p + 290.6$$

where p is the target percentage varying from 1 to 99 and the results are in millimetres.

3.2 DUM PERS

Before the fairing it is important to design everything it will contain. First of all, the frame connecting all of her velomobile parts. It could also be directly merged with the real fairing, in a monocoque structure, but with the drawback of an increased repair cost in case of accident. About other velomobile parts most of them directly come from classic or recumbent bicycles, just like wheels, brakes and transmission. Some manufacturers decided to use wheels with a larger tread than classic bicycles to improve vibration damping, but to reach this goal efficiently a dedicated damping system is recommended. Almost all velomobiles in the market come with independent dampers for each wheel, most of the times derived by automotive solutions. Dumpers are very important for urban routes. However, the limited diffusion of velomobile compatible dampers and the absence of literature references about urban load cycles do not help to easily choose the best ones. Load histories for light vehicles are an open research field. In addition, the dumping system should be designed with a geometry that is neutral with respect to the chain traction forces in order to avoid dissipation of the human power. In some cases, a damper clamping switch can be useful in order to assure the maximum stiffness in climbing. Finally, dumper stiffness has to be enough to limit the pitch movement during braking.

3.3 B RAKES

Braking system is commonly made through friction pads, drum or disc brakes. Generally, the braking system has to be more powerful because a velomobile is both heavier and quicker than upright bicycles. For this reason drum or disc brakes are more suitable. The three wheels allow a better braking distribution, dividing the effort over each one, also if the front wheels have a higher breaking efficiency due to the weight balance during decelerations. The very low position of the centre of mass makes it harder for a velomobile to overturning during hard emergency braking, a common problem for upright bicycles. If the fairing covers the wheels it is important to grant for each wheel to do not overheat, thinking they receive more than enough ventilation.

To help to choose and design the better braking system it is important to study the longitudinal dynamics of the

velomobile, as already done for automobiles [5], and the vehicle can be approximated as shown in Figure 6.

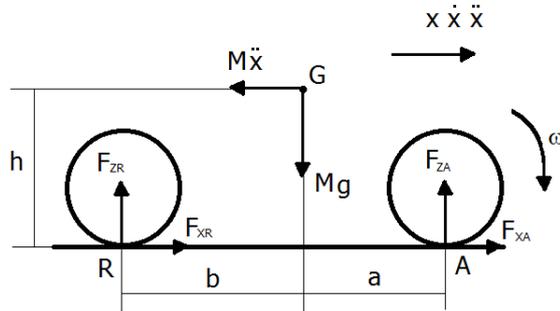


Figure 6 Velomobile forces in longitudinal motion direction.

The direction, speed and acceleration of the velomobile are represented by the x coordinate, identifying the front wheel A and the rear wheel R. The centre of masses G is located with a , b and h dimensions, while M is the velomobile weight and g the gravity acceleration.

From the balance equations of the forces and momentum it is possible to obtain:

$$F_{ZR} = Mg \frac{a}{l} + M\ddot{x} \frac{h}{l} \quad (2)$$

In Figure 7 the vertical reaction of the ground against the rear wheel is represented, varying the velomobile acceleration.

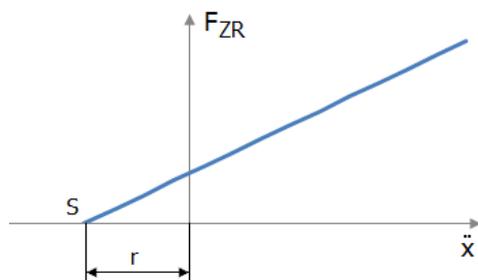


Figure 7 Vertical rear reaction force against velomobile acceleration.

Decelerations higher than point S cause a pitch movement that overturn the velomobile, so the brakes have to be designed to work in the field identified as r in Figure 7. Deceleration is linked with safety, and higher values ensure smaller braking spaces. To increase the deceleration, a good vehicle setup is needed. Some simple consideration may be derived. To increase the deceleration

$$\ddot{x} = -g \frac{a}{h} \quad (3)$$

it is a good practice to move the centre of mass in the rear down velomobile position.

By imposing the ground friction with the wheel, it is possible to link the vertical and the horizontal force components as:

$$F_{X,max} = \mu F_Z \quad (4)$$

Replacing (4) in (3), it is possible to retrieve the relation between the front and rear braking influence over deceleration:

$$\frac{F_{XA,max}}{F_{XR,max}} = \frac{b + \mu h}{a - \mu h} \quad (5)$$

The optimal braking system works on a line defined by Equation (5), as shown in Figure 8; on the same figure, constant deceleration curves are drawn.

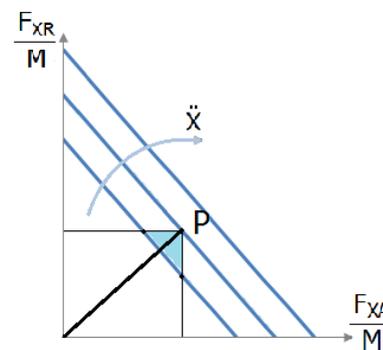


Figure 8 Vertical rear reaction force against velomobile acceleration.

Point P is the maximum braking point the system wheels-terrain allows, identifying the maximum braking deceleration possible. This value has to be compared with the point S in Figure 7, defining the closer constraint. Normally due to the general velomobile layout the more stringent constraint is the last one. Assuming this hypothesis, the braking efficiency ϵ is introduced to calculate the minimum value acceptable for the maximum braking deceleration with the chosen brakes, multiplying it with the maximum possible deceleration as in point P. This operation defines an area, cyan in Figure 8, in which the braking system has to work and to be designed: its braking characteristic has to enter this area to be acceptable.

From the maximum deceleration and the maximum vehicle speed calculated through reference pedalling cadence and transmission ratio, it is possible to define the breaking time t_f . The power of the braking system is then calculated as:

$$W = \frac{M\dot{x}^2}{2t_f} \quad (6)$$

Brakes have to dissipate the whole kinetic energy of the velomobile in thermal energy through friction contacts. Dividing this energy with the time needed to completely stop the vehicle, the brakes power is calculable.

To complete the brakes design, from the previous information is possible to calculate the pads, discs or drums dimensions and the maximum work temperature, related with the fairing design as shown later.

A velomobile also needs a parking brake that can be realized through special brake levers having a brake-holder switch (already on the market) or simply through a Velcro® strip acting on a standard brake lever.

3.4 STEERING

Velomobile steering commands are quite different from classic bicycles. Despite they keep it in the front wheels, the action is transmitted from the driver to the wheels through handlebars directly derived by recumbent bicycles or dedicated joysticks. The first system is easier to setup and cheaper than the second one, but the handlebar position could create problems in boarding movements and it has to be comfortable, without strain the driver hands. Joysticks allow an easy access to the cockpit and a comfortable driving position, but a surely more complex design than handlebars: this does not change the fact that they are wide diffuse through available trikes.

3.5 TRANSMISSION

The velomobiles transmission is usually not different from recumbent cycles, meaning that it is also directly derived from classic bicycles. It is a classic chain-ring transmission with gearshift controlled by a derailleur. Most of the times the drive wheel is the back one. This needs to introduce auxiliary idlers to transfer movement from the bottom bracket, with a longer chain-ring.

The use of a front wheel transmission would introduce different problems due to the coupled wheels steering. First of all the transmission mechanism needs to split the movement between two wheels. Also for two coupled wheels, to transmit the movement to the ground a differential gear is needed to allow wheels rotation at different speeds during corners, complicating the whole system. The last evolution of gearshift is mainly in two different directions. One is a solution with an internal hub planetary gear, improving efficiency and reliability, avoiding chain drops and lowering maintenance. This solution also allows a higher number of gears with proportional increasing of the transmission ratio for every kind of utilization. In the same time planetary gearshifts do not allow modifications by the user, an important factor when buyers are expert fans from a market like velomobiles. The other family is represented by internal hub step-less gearshifts. This solution allows using any transmission ratio inside a range chosen by the manufacturer. They are simple and versatile, with good energy efficiency, but they have limits in the maximum transmissible torque because their working concept is based on friction contacts, so they have to be designed to suit the velomobile power needs. A velomobile peculiarity in transmission design is the reverse gear, really important to handle it, especially in city. The easier solution is to provide the cockpit with two openings for the driver legs,

allowing himself to move the velomobile as needed. With the target to improve cockpit comfort, this solution is quite limiting: it is better to introduce something in the transmission to reverse the motion. It is possible to lock the free-wheel or to add a gear that user can engage manually to reverse the transmission output. The first example is easier to design but it is not safe for the final user if he forgets to unlock the free-wheel in forward direction.

3.6 PEDALLING ASSISTANCE

It is interesting to evaluate from the beginning of the project if it will have electric assistance or not. This is a primary feature for big velomobiles but could also be useful to further reduce the cyclist effort. Since the pedalling assistance needs an electric motor, it is important to verify the international regulations to know how authorities will classify the vehicle. European Directive [6] establishes that the same is not applicable to cycles with pedal assistance equipped with an auxiliary electric motor having a maximum continuous rated power of 0.25 kW until the maximum speed of 25 km/h or if the cyclist stops pedalling. For vehicles exceeding this limits, two categories are defined: "low performances mopeds", if the motor power is lower than 1 kW and maximum speed of 25 km/h, and "conventional mopeds", which could reach higher speeds. In both cases, type testing from competent authorities is compulsory but with a different level of deepening. In all member states moped classification brings along compulsory wear of a helmet, insurance and age limit and, in some case, also a number plate for the vehicles and a driving license. Technically, the pedal assistance is made through an electric motor, a control system and a battery. Since this last one is needed as energy source, it is easier to choose a DC electric motor because it does not need an inverter. The electric motor has to be linked with the main pedal transmission avoiding interferences between the two power sources. It is possible to use a planetary gear as coupling system or also to have two different transmissions for each power source, each one with its free-wheel to be synchronized together. Lead acid batteries for automotive purpose could be used to store electric energy, being a cheap solution. However, lately lithium-ion batteries are wide used because they are lighter, more durable and with higher amplitude, though they are more expensive. The control system is closely related to the cyclist pedalling movement. Generally the motor action is decided by this system on pedalling cadence and, in advanced controls, by vehicle speed feedback. If the cyclist increases his pedalling cadence, he would ask the vehicle an acceleration. By measuring these two speeds it is possible to know if the cyclist wants to accelerate and to control the electric motor accordingly.

An evolution for the control systems could be to introduce a torque feedback since it could be used to estimate the power, true expression of the cyclist effort (Figure 9). Torque sensors are the latest market news, simple and reliable, but not economic.

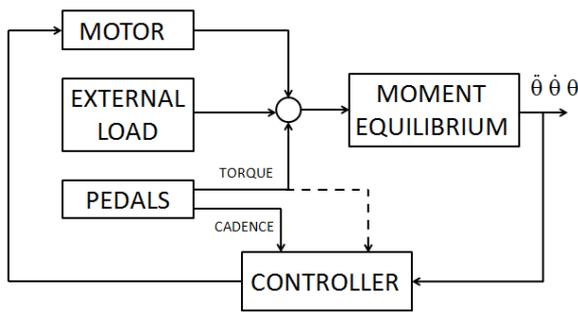


Figure 9 Generic pedalling assistance scheme.

For the pedalling assistance is also important to think about an easy way to recharge the batteries. This is always done giving the possibility to connect them to the electricity network. It is important to notice the needs of an AC /DC converter, but it is not different from the standard rechargeable battery systems from all their other possible applications. Another possibility to accompany the system is to cover the fairing with solar panels. The general non-planar shape of the fairing is not optimal for this purpose, but anyway it could be used to give extra autonomy to the velomobile or, at least, to provide energy for lights, indicators and modern accessories (i.e. GPS navigator, mp3 reader, etc.). This may increase the cost of the project considerably, so it is strongly needed to adequately calculate the corresponding benefits.

3.7 FAIRING AND LIVEABILITY

To completely define the velomobile volume around which to design the fairing, a baggage box could be introduced. It is useful for everyday needs and it could also be converted as space for a passenger, like a child, if properly designed. If the baggage box is a primary design feature, geometry and volume have to be decided in advance otherwise it is possible to fit it in the free space left by the fairing. The fairing modelling could be based on two needs: performances or practicality. For urban purpose performances are not as much important as practicality, then fairing could just be designed as a driver protection. Basic aerodynamics knowledge is more than enough also because an exasperated aerodynamic shape is far to fulfil comfort requirements. NACA curves [7] can be used as references to design the fairing, because they are easy to use and grant the lowest possible aerodynamic resistance. The cockpit ventilation is a subject of primary importance for the cyclist comfort. The velomobile must be designed to ensure oxygenation to the driver during the pedalling effort. Through vents is possible to remove the heat produced by the cyclist and also the water vapour, in order to keep controlled the air inside the closed space of the cockpit. Vents are designed by the air flow they have to grant, calculable from an energy balance, as shown in Figure 10 where E is the external ambient, I the cockpit volume (at temperatures T_E and T_I respectively), \dot{m}_v the water vapour produced by the cyclist (sweat plus breathing), \dot{m}_A

the air flow from the vents, \dot{Q}_S the thermal power from the external (due to solar radiation and convection with the ambient air), and \dot{Q}_V the thermal power produced by the cyclist during the physical effort.

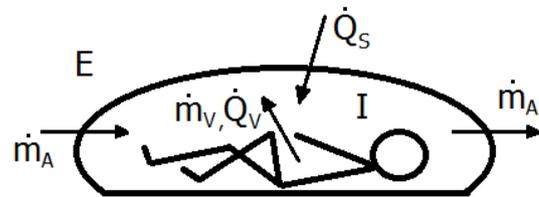


Figure 10 Cockpit energy balance.

The thermal balance is:

$$\dot{Q}_S + \dot{m}_v C_{pH_2O} (T_I - T_{ref}) + \dot{Q}_V + \dot{m}_A C_{pA} (T_E - T_I) = 0 \quad (7)$$

where T_{ref} is the reference human temperature, usually around $37^\circ C$ but it may be higher during physical activity, and C_p is the mass heat capacity of the water and the air respectively.

By fixing a reasonable internal temperature $T_I = 24^\circ C$, external summer condition $T_E = 30^\circ C$, by estimating the thermal power from the external through the velomobile surface, the thermal power produced by the cyclist and the water vapour flow, it is possible to calculate the air flow used to design the vents. General values for the human vapour flow may vary from 4 g/h to 330 g/h in normal conditions up to 1000-2000 g/h for a very intense physical effort.

To design the vents it is possible to use the default NACA profile [8] as shown in Figure 11. It consists of a submerged-duct entrance taking advantage of the turbulence at its edges to generate a depression that draws air inside the cockpit without interfere with the fairing external air flow.

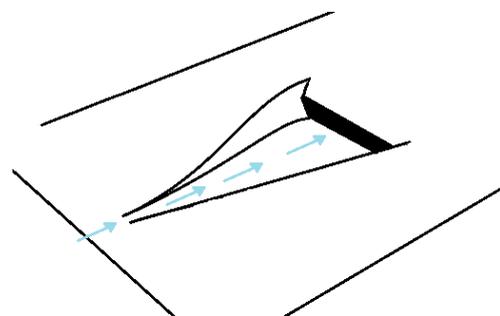


Figure 11 Sketch of submerged-duct entrance.

The dimensions are calculated through graphs and tables from the dynamic pressure of the air flow on the fairing where the duct entrance is located (determined by CFD analysis), the differential pressure between inside and

outside the cockpit volume (considered equal to 1 for this purpose), the air flow speed and the inclination of the duct entrance ramp.

The flow speed is related to the previous calculated air flow through the air density and the duct section area:

$$S = \frac{\dot{m}_A}{\rho_A V_A} \quad (8)$$

It is also important to design out-takes air duct, otherwise the cockpit would be under pressure by the entering air flow.

Normally the wheels openings are more than enough, but with the idea to improve the driver comfort may be better to

insulate the cockpit from wheels, resulting in needs of such out-takes ducts. Defined the fairing shape, the next step is to design the access door to allow the driver to enter the cockpit. For each project size the reference user is the one with the biggest dimensions. Anywhere the lock mechanism is really important and it has to be simple and resistant for an everyday utilization. Some different types of openings exist, each one with strengths and weaknesses points. Hinges or slides can realize the movement of the door as shown in Figure 12. This last system occupies less space but it is more complex to build, while hinges are a good compromise between simplicity, space and structural stiffness.

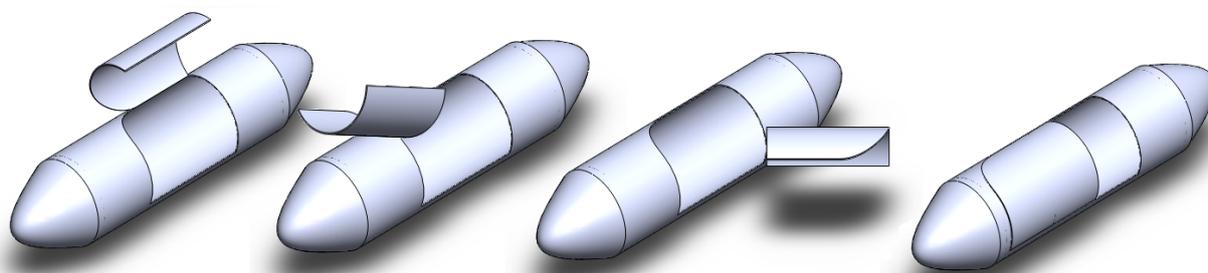


Figure 12 Different types of access door.

Together with the fairing, seat regulations have to be added allowing each user to adapt the velomobile to his characteristics. These two steps are linked because regulations may create interference between the cyclist and the fairing when he modifies his position in the cockpit. Seat regulations may be done through two ways: varying the seat or the pedal position. Usually large regulations are left to the pedal position while little regulations are made through the seat position, because this last type is more complex to design due to the seat links with the frame.

About the fairing construction two alternatives are available: a frame covered by a light fairing or a monocoque where the frame and the fairing are merged in a single element. The first solution is easy to design and build but also leaves little optimization margins. If compared with the monocoque solution, to reach the same mechanical strength the vehicle is heavier, affecting the performances. Monocoques are generally lighter and stiffer but in some cases the design may bring to unbui ldable, not practical results. Finally the better solution is generally well matched with composite materials, lighter than metals, stiffer than plastics and a lot more mouldable than other solutions.

While normally more attention is reserved to the exterior side of the fairing, cockpit liveability keep increasing its importance through manufacturers: acoustic and thermal insulation, materials and detailing aim to improve the driver comfort.

3.8 WINDSHIELD

After the fairing is completely defined, windshield has to be adapted to it. The main objectives of the windshield are to protect the cyclist and to ensure a good field of view around him. Some velomobiles do not have a windshield; others have it fixed while solutions with a removable one exist too. These last cases are more complex to design because they need removable links but they solve cyclist ventilation problems, especially during the warmest period of the year. Anyway it is crucial to design an intakes air system to feed the cyclist needs and to remove humidity formed by the sweat of the pedalling effort. Back to the driver field of view, no directive to design it exists at present. The only requirement is to have a good forward and lateral vision, while the backward vision is granted by rear-view mirrors. As guideline to define the windshield and the cyclist field of view, automotive Directives D77/649 and D88/366 for M-type vehicles may be used [9]. They give the needed information to setup a vision pyramid around the driver. This geometry is then intersected with the fairing, showing the area where the windshield has to be transparent. Visual fields obtained with this method applied to the velomobile driver typical lying position produce a shadow zone in the front-down direction of the field of view, just like in front of automobiles forward bumper. Therefore this method may be used for the first approach, improving the field of view accordingly.

Lastly, even if not prescribed by law, lights, indicators, horn, reflectors and eventually safety belts are all good options to include in a commercial project, for driver safety.

4 CONCLUSIONS

The paper sets a design guideline for velomobiles, covering the main topics that characterize these vehicles. General information are given and a lot of different possibilities are left to be chosen by the manufacturer.

A unique solution does not exist and each choice has its strengths and weaknesses. As the discipline of velomobile design is quite new, each topic discussed in the present paper needs to be further expanded in details through specific research in order to make the velomobile an efficient and safe option for a green personal mobility in the urban context. By using the presented information it is then possible to know the basis to start a new velomobile project and to plan the needed work.

REFERENCES

- [1] <http://www.recumbentjournal.com/news/industry/iteam/533-recumbents-on-the-rise-according-to-leisure-trends-group.html>, [24/05/2012].
- [2] Chen G., Kautz S.A. and Zajac F.E., Simulation analysis of muscle activity changes with altered body orientations during pedaling. *Journal of Biomechanics*, Vol. 34, No. 6, pp. 749-756, 2001.
- [3] Too D. and Landwer G.E., Maximizing performance in human powered vehicles: a literature review and directions for future research, *Human Power Journal*, Article No. 16, 2008.
- [4] Reiser II R. F., Peterson M. L. and Broker J.P., Anaerobic cycling power output with variations in recumbent body configuration, *Journal of Applied Biomechanics*, Vol. 17, No. 3, pp. 204-216, 2001.
- [5] Guggiani M., *Dinamica del veicolo*, Città Studi, 2007.
- [6] European Commission, Directive 2002/24/EC relating to the type-approval of two or three-wheel motor, 2002.
- [7] Abbott I.H. and Von Doenhoff A.E., *Theory of wing sections*, Courier Dover Publications, 1959.
- [8] Frick C.W., Davis W.F., Randall L. and Mossman E.A., *An experimental investigation of NACA submerged-duct entrances*, Patent NACA-MR-A5E23, 1945.
- [9] Morello L., Rossi Rossini L., Pia G. and Tognoli A., *The automotive body volume II: system design*, Springer, 2010.
- [10] Wilson D.G., *Bicycling science*, The MIT Press, 2004.

A METHODOLOGY TO DESIGN ROBOTIC ARMS FOR SERVICE TASKS SINCE EARLY DESIGN STAGE

C. Castejón* G. Carbone** J.C. García-Prada* M. Ceccarelli**

* MAQLAB group, Dep. Ingeniería Mecánica, Universidad Carlos III de Madrid. Av. De la Universidad 30, Leganés, 28911 Madrid, España, {castejon; jcgprada}@ing.uc3m.es

** LARM: Laboratory of Robotics and Mechatronics, DiMSAT, University of Cassino Via Di Biasio 43, 03043 Cassino (Fr), Italy, {carbone; ceccarelli}@unicas.it

ABSTRACT

In this manuscript, a methodology is presented for designing robotic arms for service tasks. This methodology attempts to take into account analytically the specific novel requirements that are needed for service tasks. This goal is achieved by means of the formulation of specific objective functions which can evaluate and compare the quality of a design among others. Additionally, these objective functions are investigated for being implemented in suitable optimal design algorithms. The proposed methodology can be extended to any robotic mechanism for service tasks. The design method begins with the development of type and number synthesis to identify the minimum needed number of degrees of freedom. After selecting the most suitable mechanism, the proposed design procedure uses the objective functions to compare the most acceptable solutions. Finally, by means of a numerical algorithm, the dimensional synthesis is performed to obtain optimal length and thickness of each link. The proposed procedure has been applied for a case of study in order to prove its feasibility and practical usefulness.

Keywords: service robots, robotics design, multi-objective optimization, robotic arm

1 INTRODUCTION

New fields of robotics applications are characterized by specific requirements that existing robot solutions may not be able to fulfill. Therefore specific investigation and designs activities are required (Ceccarelli, 2009). Currently, the new trends are demanding novel designs which consider new requirements such as: safety, low cost (both, economic and energetic), optimal ratio weight/load, morphological resemblance with humans (in order to enhance acceptance and integration), and the simplicity in

the movement control related with the number of degrees of acceptance and integration), and the simplicity in the movement control related with the number of degrees of freedom (DOFs) (IFR, 2011). For this reason, there is an increasing interest in new actuators and mechanism designs (Wu et al., 2009). In the literature some optimal design approximations can be found in order to enhance different aspects such as safety for humans and robots (Gao et al., 2000; Park et al., 2010), movement precision with the minimum links weight (Carbone et al. 2003) or topology (Logo et al., 2009). Synthesis of a mechanism can be defined as the process to develop and select the configurations, materials, geometries and dimensions of a mechanism. In industrial design, synthesis is usually missing due to the huge experience designers have in a specific mechanism they usually design as based on specific requirements. Normally, the process consists on the enhancement of a previous design.

Contact authors: Cristina Castejon¹, Giuseppe Carbone²

¹ Av. de la Universidad 30, 28911 Leganés, Spain
E-mail: castejon@ing.uc3m.es

² Via Di Biasio 43, 03043 Cassino (Fr), Italy
E-mail: carbone@unicas.it

Nevertheless, synthesis is a key issue to generate the needed know-how not only when previous designs are not available (Johnson, 1978).

In the synthesis process, three stages can be identified: type, number and size synthesis. In general, there are many solutions to the design process and infinite solutions for the dimension and materials selection.

Thus, computer aided optimization processes can be very useful at this stage. The close interaction between robots and humans makes not suitable the classical robot design methods applied to industrial robots. In fact, interaction between robots and humans and safety became two of the most important issues (Park et al., 2010), while the need to achieve a high payload or a high speed ability (basic requirements in industrial robots) is reduced. Those aspects must be considered at the first stages of the design process (Castejon et al., 2007). Some researchers suggest that the robot design would be optimized in the sense of increasing the dexterity in the human workspace (Yan, 1998). On the other hand, light weight robot structure is advisable in order to prevent damage to humans and own robot integrity. In this case, the optimization of the stiffness with the aim to enhance the precision and light weight design is needed, as is justified in (Bai, 2010). Power consumption and path planning can be also considered to enhance the autonomy and flexibility as is proposed in (Blanco et al., 2005; Font-Llagunes et al., 2009). The optimization related to any of the criteria commented before produce a final design that is not optimal with respect to other criteria. As a possible solution, in this manuscript a multiobjective optimization is proposed. The design criteria are stated in an analytic way and it is related to the task the robot is going to perform. Some authors perform optimization in the design process (Lee et al, 2010) of a manipulator; in terms of dexterity, stiffness and workspace.

Despite the multiobjective interest, authors calculate a global objective function which includes the three subjects of interest by means of weighted normalized index. Dimensional synthesis is performed in (Gao et al., 2010) in order to discover the optimal geometric configuration according to objective functions and geometric restrictions by means of performance indices based on stiffness and dexterity. The solution of the optimization problem is achieved by means of Genetic Algorithm. Design indices are also used in (Kim et al 2009), which perform a kinematic optimal design in order to obtain a parallel mechanism. The indices considered in this work are workspace size and global kinematic isotropic index of the mechanism. They obtain a plot which relates the size of links but do not obtain a unique solution.

The same philosophy are used in (Yoon et al 2010) for parallel mechanism. In this work, a design methodology to build non-conventional robotic structures since their first stages is presented.

To achieve this goal, a methodology proposes the use of a link synthesis inside the creative process (Shigley et al.,

2003), where all the non interesting characteristics are a priori erased in the first design.

The characteristics to work in the first stage will be: type of link, distance between joints and type of joints (Merlet, 2005). On the other hand, to optimize a robot structure with respect to a single design criterion produces a non optimal design with respect to other design criteria. A possible solution to solve this limitation, proposed in this manuscript, is the implementation of a multi-objective optimization process, which works with different criteria at the same time, as proposed in (Carbone et al, 2007; Castejon et al, 2010).

2 DESIGN METHODOLOGY

Different flowcharts about creative design process can be found in literature [8, 10], but all of them are not centred in the service robot design and they are oriented for production purpose. The creative design methodology proposed in [10] is based in the generalization of previously existing kinematic chains and the development of atlas which include a set of specialized chains and designs. The flowchart starts with the identification of existing designs which fulfil the design requirements in order to define the topological characteristics.

For this reason, in the flowchart proposed by this author there is no place for creativity in new tasks as service robots demands. On the other hand, there is no optimization process (essential in service robotic design, where constraints can offer contradictory solutions) and feedbacks between steps. The mechanical design process proposed in [8] let space to creativity in the mechanical design synthesis steps and optimization is proposed in the topological and dimensional synthesis steps. Nevertheless the process proposed by Johnson is manufacturing oriented, we mean, the aim is to come in the production-distribution-consumption-recovery cycle and new requirements are not considered in the mechanical design process.

The complete design process proposed in this manuscript includes optimization process in the dimensional synthesis where particular specifications for service robots are included. This process can be divided in different fundamental parts that constitute its morphology. Figure 1 shows the relations among different components. The process starts with the definition of the problem where, data and constraints are stated based on the task the robot must perform. On the other hand, the ultimate design goal is obtain a service robot that fulfil the task specified and with the minimal cost (energetic and economic).

The sequential steps of the process are the following:

- 1) Give data & constraints: The first basic part of the design is the acknowledgement and definition of the problem. This step allows setting the specifications the final prototype must fulfil. We understand as specification the detailed precise statement describing a product (Johnson, 1978).

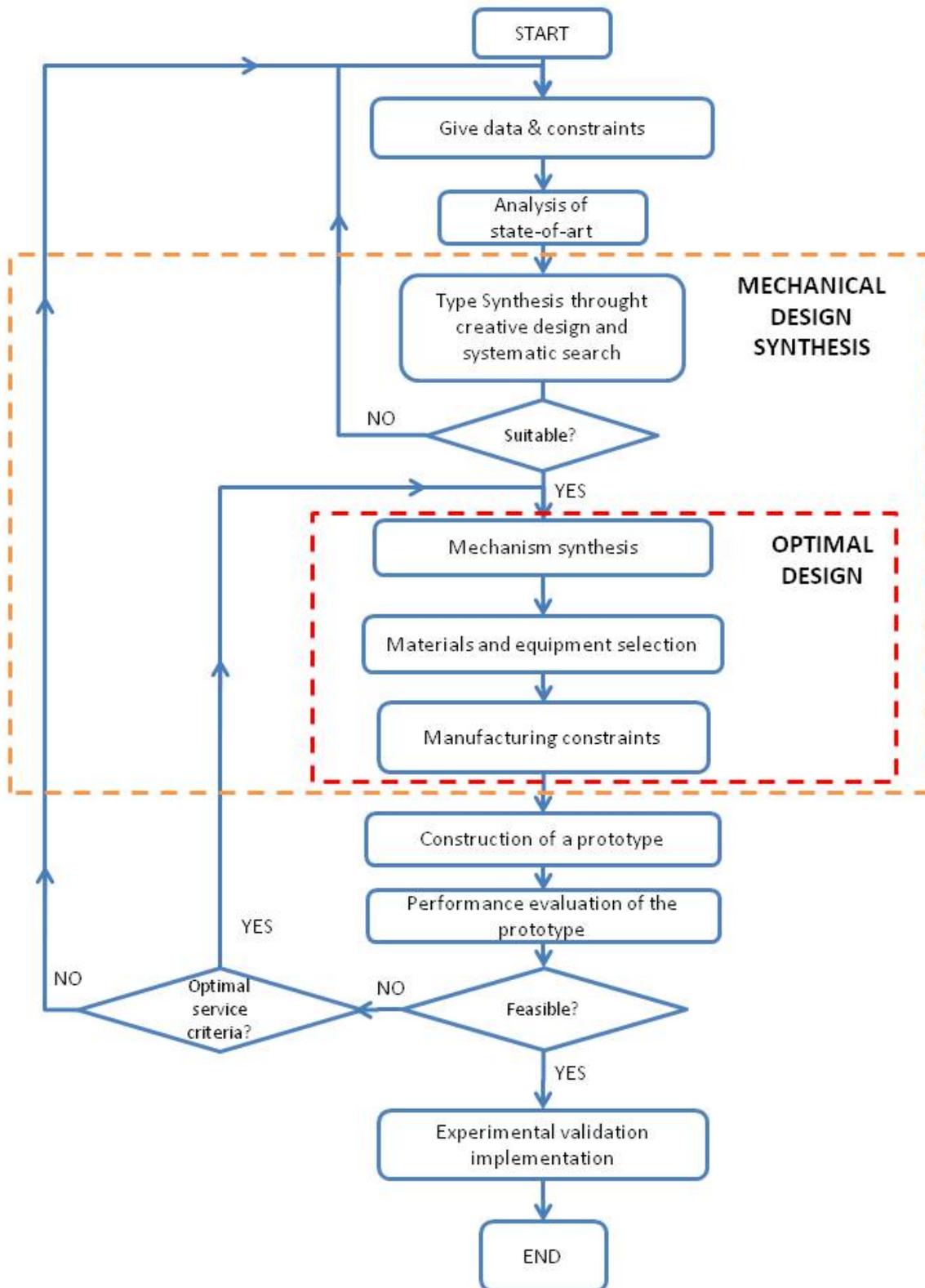


Figure 1 Design process methodology.

In this particular case, specifications are related with the service task and they will meet technical aspects (kinematic, dynamic, structural point of view) and no-technical aspects (human adaptability, aesthetics, etc.). On the other hand, constraints must be also detected in this step, in order to define completely the design problem.

- 2) Analysis of state-of-art: In the next step, the compilation of the information about the existing solutions for similar problems is accomplished (based on the information obtained from previous step). Then, the search of information about similar solutions, patents or designs developed must be done in order to evaluate the different possibilities designers have.
- 3) After collecting information and stated the most important characteristics and constraints of the service robot, the type synthesis must be developed. Type synthesis step is based on the search of the basic kinematic chains which they can, a priori, fulfil the specific requirements in the problem definition. This part needs generally an important creative effort besides the correct application of the systematic techniques of synthesis [10]. Number synthesis establish a set of basic links with a specific number of degrees of freedom (DOF) and links transformation techniques will give more solutions for the same problem. In general, some solutions will be found and the most promising mechanism must be selected. This selection is done by comparing different characteristics of the mechanism, such as the workspace or some kinematics aspects. If the kinematical chain selected can fulfil the specifications and constraints previously stated, then, the design process can continue with the optimal design process. If this statement is not true, then a reconsideration of the specifications defined must be done.
- 4) With the most interesting kinematic chain selected, mechanism synthesis is performed in order to select the dimensions of the structure. This synthesis is carried out based on the optimization, by means of the selection of criteria which define the most important requirements the robot must fulfil. In the most complex cases, optimization techniques must be applied; meanwhile the easiest cases trial and error techniques with some easy calculus will be enough. For a service robotic arm, different and contradictory aspects are taken into account since the beginning of the design process [21]. For this reason, in this work a multi-objective design approach is proposed in order to take into account several design criteria at the same time.
- 5) With the optimal dimensions of the robot structure, the selection of the suitable materials and associate equipment (such as sensors, actuators, end-effectors, etc.) must be defined. This step is preliminary to the statement of the manufacturing constraints, which

determine is a prototype, that must be constructed, is feasible or not.

- 6) In this step, and based on the results obtained in previous steps 4 and 5, manufacturing requirements must be detailed in order to complete the specifications. These requirements will be of interest for the next step related with the constructions of the first prototype.
- 7) After being completed specifications for the prototype, robot must be constructed with all the measurements, equipment and constrained stated.
- 8) In the Performance evaluation prototype step, the robot will be tested and evaluated in order to guarantee the kinematical and dynamical characteristics the robot must fulfil. In this step theoretical calculus are compared with some basic experimental results. If the performance evaluation is not positive, then, the optimal service criteria are evaluated to know the origin of the problem, which can be dimensional (in this case a feedback to the first step of the optimal design must be done) or is the selection of the basic specifications, so, in this case a feedback to the first process step will be done.
- 9) If the performance evaluation is positive then, experimental analysis will be the real proofs of the success or failure of the design. After tested the success of the product, it will be ok to enter in the production, distribution, consumption and recycling cycle, which is the real objective of a product.

The main part of this work is the step related with the optimal design process (points 4, 5 and 6 of the methodology proposed), and particularly point 4 which makes reference to the dimensional synthesis. The multiobjective optimization technique for obtaining the links dimensions is developed in this manuscript. To accomplish this goal with success, the specifications and constraints related with service robot tasks will make up the formulation of the optimization problem. Besides, the application of the Mechanical design synthesis process will be done for a case of study in order to test the goodness of the methodology proposed.

3 OPTIMIZATION PROCEDURE

In the dimensional synthesis procedure, we start on the not optimal configuration obtained from the type and numerical synthesis. Optimization is the cornerstone in the dimensional synthesis process. In this work, the obtaining of the optimal design in different aspects, or criteria are looking for, so the common optimization techniques do not completely cover the problem and we must focus the study in a more specific and complex optimization part which is the multi-objective optimization. Optimization problem refers to choosing the best solution from some set of available alternatives. Different searching algorithms can be used in order to reach the best solution.

The algorithms can be divided in two groups: stochastic or iterative methods, such as the gradient descent criteria [22], or heuristic methods, recommended for non linear and complex problems. One of the most popular is the method based on Genetic Algorithms [23]. The definition of the problem and the complexity of the objective function will determine the best selection of the searching algorithm. In this work different criteria are needed to be included in the optimization problem, so a multiobjective optimization problem must be solved. The searching algorithm used is the gradient descent which, in our case, has been proved to give good results. Optimization criteria F_1, F_2, \dots, F_k , are functions of the design variables (x_j) which help to the evaluation of the solution quality in form of objective functions. So we can formulate the problem of optimization in a mathematical way as in Equation 1.

$$\text{Min}(F(x_j)) \text{ where } F(x_j) = [F_1(x_j); F_2(x_j); \dots; F_k(x_j)] \quad (1)$$

By using the solution of the problem by means of the multiobjective optimization, which tries to find, in each iteration, the solution that minimize the function with the maximum value, the search algorithm is based on the descent gradient. Efficient solutions will be obtained, and no a unique solution as in other optimization problems. Local and global minimum values are one of the problems found in these algorithms, especially when we work with non linear objectives functions. These efficient solutions produce an optimal result of one, or several optimization criteria. We must take into account that, when several criteria are stated in order to minimize in a simultaneous way, probably a variety of different feasible and efficient solutions will be existed.

3.1 OPTIMIZATION CRITERIA

Optimization criteria define the objective functions for the optimization problem. The same criterion can be formulated with different analytical expression. The selection of the expression that will define the objective functions will affect to different aspects of searching solution, such as the convergence to a unique solution or the time computing. In this section the analytical expressions selected for the optimization process are enumerated.

3.1.1 Reach function

This objective function tries to achieve the major distance between the end-effector and the robot base. The reach calculus are performed in the point called as extended point (X_{ext}, Y_{ext}) and defined in Equation 2.

$$f_1(x) = A = \sqrt{X_{ext}^2 + Y_{ext}^2} \quad (2)$$

fminimax of MATLAB® toolbox [24] used in the optimization program always tries to minimize the components of the objective function vector (see Eq. 1). So, component $f_1(x)$ must decrease when the reach increase. On the other hand, for a better performance of such a function all the components of the objective function vector

must have similar values when the mechanism is closed to the optimal values. For that reason, we look for the components of the vector $F(x)$ which have values near to zero when they are optimum.

We mean, a normalization of the criteria must be done in order to compare.

3.1.2 Weight function

The function calculates the mass of the links, by means of the volume and density of them as Equation 3 shows:

$$f_2(x) = P = \sum_{i=1}^8 l_i \cdot e \cdot w \cdot \rho \quad (3)$$

Where, l_i are the link lengths, e is the thickness and w is the wide of the link i .

3.1.3 Movement amplitude

The objective of this function is avoid to achieve solutions in the optimization process where the trajectory of the end-effector has a small traveled, we mean, the points called retract (X_{ret}, Y_{ret}) and extended are too close, reducing the workspace of the robot arm. This criterion searches higher amplitude in the trajectory by means of Equation 4:

$$f_3(x) = AM = \sqrt{(X_{ext} - X_{ret})^2 + (Y_{ext} - Y_{ret})^2} \quad (4)$$

The third component of the objective function will have a small value the more increase the movement amplitude of the mechanism.

3.1.4 Shape level

This criterion is focused on modeling the lower shape of the trajectory with the aim of find a soft trajectory. The objective is to limit the appearance of sharp changes in the trajectory that will produce huge accelerations in the movement with the consequents inertia efforts. This function is very interesting for manipulators with *pick and place* task such as the example presented in this work. The shape level is calculated from the link lengths (L_i), and evaluates the significant points of the trajectory, which are highlighted in Figure 2.

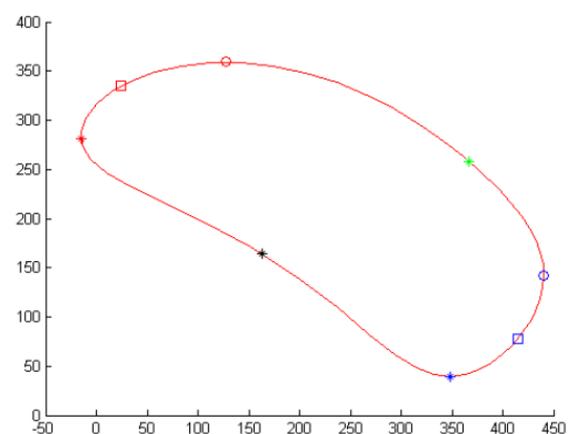


Figure 2 Significant points of the trajectory.

The shape level lets measure the separation between a part of the mechanism trajectory and its interpolation with a parabola built with the significant points.

By calculating this difference, the resemblance degree between the trajectory and a parabola or a third degree soft function is estimated. The comparison with the parabolas is performed in two parts of the trajectory, the lower part, between the minimum and maximum X-coordinate, and on the other hand the upper part between the same points. This parameter is used to test that the parts of the trajectory which correspond to the maximum extension and the area around it have soft shape and a related resemblance with the interpolation functions. Furthermore, the interpolation benefits the trajectories were elongated in the vertical direction, so we can fit the shape to the task proposed in the application. Formulation of $f_4(x)$ is presented in Equation 5:

$$f_4(x) = \left| 1 - \frac{N_j}{N_0} \right| \quad (5)$$

Where, N_0 is shape level of the initial configuration and N_j is the solution obtained for each iteration. The shape level is considered suitable when values are in range [0,10]. If the values are greater than 100, we will be considered the trajectory do not fulfill the demanded characteristics.

3.1.5 Extended point height

The last component of the function $F(x)$ tries to close the extended point of the mechanism trajectory to the possible lowest point. For the application of study (see Figure 4-h) that position is obtained when link AC is in the vertical position and point C is under the link AB. To achieve this position, optimization program will calculate $f_5(x)$ by means of Equation 6:

$$f_5(x) = Y_{ext} + \overline{AC} \quad (6)$$

Of course, more criteria can be added to the optimization process such as safety, stiffness, task properties, etc.. In this work, previous criteria represent the most suitable for a simple manipulator prototype. They have been included as an example to be applied to test the soundness of this proposal.

4 A CASE OF STUDY: A ROBOTIC ARM DESIGN

As a practical application to test the proposed methodology, a robotic arm that will perform the pick and place task will be design.

4.1 CREATIVE SYNTHESIS

In the type synthesis, Equations 7 and 8 will be used to obtain the most promising kinematic chains. Equations 7 and 8 are obtained from Grübler equation [18]:

$$G = B - Q - 2P - 3 \quad (7)$$

$$T + 2Q + 3P = L - (G + 3) \quad (8)$$

Where G is the degrees of freedom required, B , T , Q and P are the number of binary, ternary, quaternary and pentagonal links respectively and L the total number of links (we mean $B+T+Q+P$). To find the simpler mechanisms we started to choose only binary links ($B \neq 0$) and after that, an increasing number of ternary links are introduced until a variety of simple mechanisms, which fulfill the constraints previously established, is achieved (in this case we try with $T=0$, $T=1$ and $T=2$).

In Figure 3 mechanisms selected after applying the associated linked concept are presented. All the mechanisms presented has a configuration $L=6$ ($B=4$, $T=2$), and are formed by complete joints. It must be noticed that the intention is not to find all the possible solutions (configurations), but only the most interesting to take an intelligent decision in order to obtain the most efficient mechanical design. With the selected mechanisms (in Figure 3) the study continues with the comparison process. Working Model 2D[®] (WM2D[®]) software has been used to this task. WM2D[®] let to evaluate the behavior of different mechanism in an easy and fast way. Selection criteria for the most promising mechanism have been the workspace, maximum acceleration and velocity, robot dimensions, the easy construction and the efforts the links support. Results obtained from this comparative study are summarized in table I.

In order to measure and compare the workspace we approximate the trajectory of the end-effector to an ellipse where L_{SM} and L_{Sm} correspond to the semi-major and semi-minor axis respectively. Maximum acceleration and velocity of the end-effector will be measure for different rotor velocities (Π , 1.57, 0.79, 0.39) and the last column of the Tab 1 represent the maximum velocity, the mechanism need to achieve in order to obtain a stated acceleration in the end-effector (that in this work is 0.5 m/s).

Based on the results presented in Table I, we consider the reach and the maximum speed ω_{max} as the most interesting parameters to compare. Reach represent the length of the links, which is directly related to the mechanism weight (that has influence on the economical and energetic cost). In this case, and considering similar reaches for all the mechanism to be compared, it will be interested to search the lower reach. On the other hand, maximum rotational speed is advisable to choose the solution with the higher value. With this discussion, and based on results of Table I mechanism 3 and 4 are the most promising.

Taking into account mechanism 4 present a one-dimensional trajectory (as we can see in Table I, $L_{Sm}=0$), it will have less interest for the multiobjective optimization due to the simplicity of its trajectory. So, the most promising mechanism that has been selected is the mechanism number 3 presented in Figure 3-c. The final configuration placed for the specific task can be seen in Figure 4h.

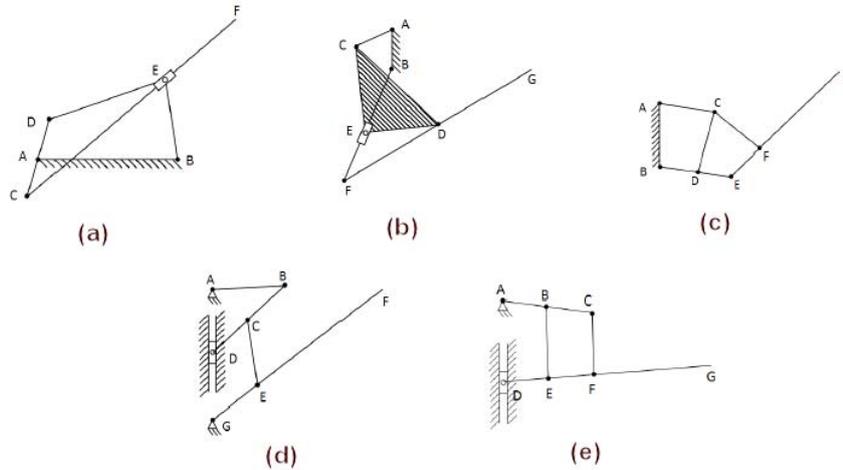


Figure 3 Selected mechanisms: (a) type 1; (b) type 2; (c) type 3; (d) type 4; (e) type 5.

Table I - Summarize of results for the mechanism solutions comparison

ID	Dimensions (mm)	L_{SM} (mm)	L_{Sm} (mm)	Reach (m)	ω (rd/s)	a_{max} (m/s ²)	V_{max} (m/s)	ω_{max} (rd/s)
1	AB=150; DC=100 BE=170; DE=170 CF=700	501.20	72.80	1.29	Π , 1.57 0.79 0.39	7.18 1.80 0.45 0.11	1.40 0.70 0.35 0.18	07.83
2	FD=165; BF=250 AB=100; FG=480 AC=60; CD=232 CE=160; DE=138	503.29	70.71	0.89	Π , 1.57 0.79 0.39	16.22 4.06 1.02 0.25	2.14 1.07 0.53 0.27	0.55
3	AB=150; CF=210 AC=80; EG=340 BE=210; BD=135 CD=170; EF=86	502.49	178.04	1.16	Π , 1.57 0.79 0.39	6.81 1.70 0.43 0.11	1.46 0.73 0.37 0.18	0.85
4	AB=100; GE=63 AG=200; CE=150 BD=130; BC=65 GF=400	502.89	0.00	0.98	Π , 1.57 0.79 0.39	14.74 3.73 0.93 0.23	1.72 0.86 0.43 0.22	0.5
5	AC=100; DG=360 AB=50; DE=103 BE=120; EF=64 CF=130	530.00	56.57	0.71	Π , 1.57 0.79 0.39	101.2 5 38.92 9.67 2.63	3.35 1.68 1.01 0.51	0.17

Table II - Objective function values

	Initial	Execution 1	Var 1 (%)	Execution 2	Var 2 (%)	Var (%)
Max. Reach	431,42 mm	492,92 mm	14,25	531,48 mm	7,82	23,19
Weight	0.44 Kg	0.34 Kg	-22.82	0.31 Kg	-7.24	-28.41
Amplitude	475.42 mm	580.80 mm	22.16	679.10 mm	16.92	42.84
Shape level	1.00	1.87	86.59	1.91	2.60	91.45
$Y_{ext} + AC$	64.83 mm	-19.52 mm	-130.11	-59.99 mm	207.25	-192.53
F_{max} retracted	35.72 N	30.03 N	-15.93	28.91 N	-3.70	-19.04
F_{max} extended	32.77 N	29.32 N	-10.52	25.00 N	-14.74	-23.71

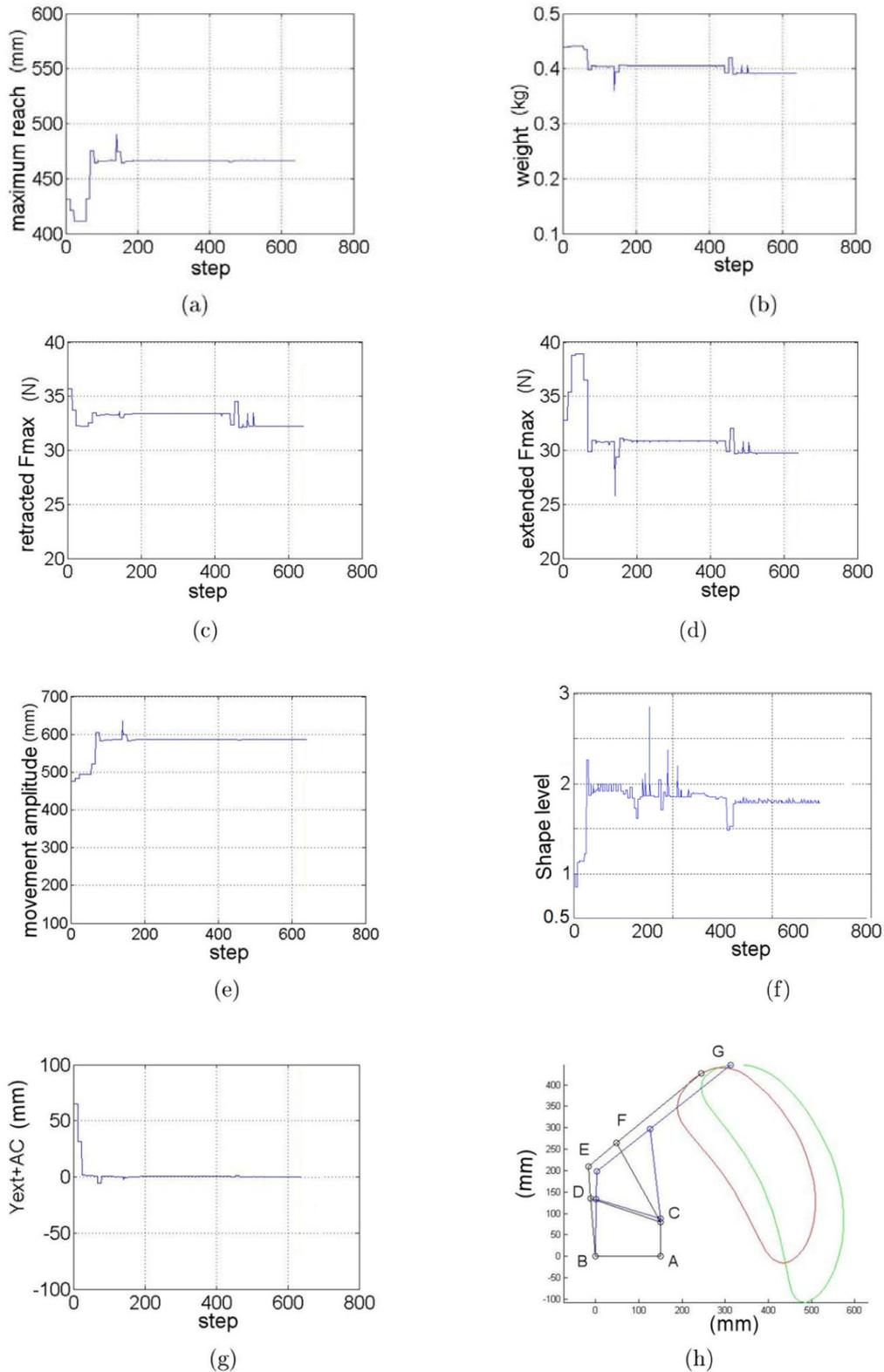


Figure 4 Results of the optimization process:

(a) Maximum height objective function (OF), (b) weight OF, (c) Maximum force with the arm retracted, (d) Maximum Force with the arm extended, (e) movement amplitude, (f) shape level, (g) extended point height, (h) initial mechanism (in black) and its trajectory (in red) vs. optimized mechanism (in blue) and its trajectory (in green).

Table III - Optimal decision variables values

variables	Initial value (mm)	Step 1 (mm)	Var 1 (%)	Step 2 (mm)	Var 2 (%)	Var 3 (%)
AB	150	151.20	0.80	160.55	6.18	7.03
AC	80	88.45	10.57	94.73	7.10	18.42
BD	135	132.83	-1.61	139.62	5.11	3.42
CD	170	155.72	-8.40	137.30	-11.83	-19.24
CF	210	210.67	0.32	224.04	6.35	6.69
DE	75	66.16	-11.79	61.90	-6.43	-17.46
EF	86	158.23	83.99	165.94	4.87	92.95
FG	254	238.14	-6.24	247.89	4.10	-2.40
thickness	7	5.22	-25.48	4.72	-9.56	-32.60

4.2 MULTIOBJECTIVE OPTIMIZATION IMPLEMENTATION

Multiobjective optimization is performed by means of the *minimax* function available in the Optimization Toolbox of MATLAB[®]. Starting with an initial solution, the algorithm calculates the gradients of the objective function in all the directions of the hyperspace of the decision variables. If the gradient is higher than an adjusted tolerance, the program calculates a new position in the suitable direction of the hyperspace.

In that way, the optimal point will be achieved in an iterative way. The program will be able to finish for several reasons: the most common because the function converges to a solution, because there do not exist gradients of the objective functions with values above the imposed tolerance, because the search direction magnitude is lower than the specified tolerance, because the maximum number of iterations is achieved or because feasible points are not found. Any of the first two reasons is an acceptable solution. Similarly, link dimensions and thickness are considered by the program as a vector \mathbf{x} with size nine, where the first eight components are the link lengths and the last component is the thickness.

The optimization problem must have constraints set in order to complete the complete approach. In this case, the main shortcoming is that the results offered by the program must correspond with a feasible mechanism which can be constructed. Therefore, another constraint as the maximum payload the mechanism can support has been considered with the application of a deformation limit. The results produced with the first execution of the program are presented in Figure 4. The Figure 4a shows the evolution of the maximum reach versus the number of iterations. It can be clearly seen how in the first iterations values decrease in order to minimize the other objective functions and in then the values increase until, in first instance, a stable value around 460 mm, and in iteration number 625 a sharp change is produced obtained the final value in 492.92 mm.

The initial decrease in the maximum reach is corresponding with the total mass decreasing, the increasing of the maximum amplitude and the decreasing of the G point height in the extended position. The previous values enhancement produce that, initially, the maximum reach decrease. Only when the maximum reach value is lower than the rest, the algorithm will try to optimize this component. On the other hand, the sharp change in the last iterations is produced by the violation of the constraints imposed to the problem. Due to the change coincide with the minimum acceptable forces values under the limit imposed (see Figure 4 c) and with a drastic increase of the shape level (over the maximum value imposed, see Figure 4 f), we can state that these criteria have caused the change. The rest of the plots presented in Figure 4 evolve in the improvement direction sense. The configuration of the initial and the optimized mechanisms and the trajectories of their end-effectors can be seen in Figure 4h. In this figure we can observe that the optimization process has changed the link lengths increasing mainly the link EG and decreasing CD link. Besides, the thickness has been reduced from the 7 mm (which was the initial value) to 5.21 mm. The program has been running one more time in order to avoid minimum local solutions. In Table II the objective function values at the end of the optimization program for each execution are presented (algorithm has been executed three times to verify the convergence). As can be seen the thickness has been reduced until the imposed limit (feasible constraint). In the same Table II columns labeled as Var 1 and 2 represent the variation percentage of each objective function component after each execution related to the initial value (in each execution). The last column *Var* represents the total percentage variation between execution 2 and the initial value (in the first column). The second execution performs a substantial enhancement of the initial requirements. All the components of the objective function achieve better results obtaining in some cases to find the imposed limits.

A third execution has been running in order to test the previous results not obtaining better results.

In Table III results of the decision variables are shown. Columns with label Var represent the same concept than in the previous Table I. As can be seen in this Table III, percentage variation of first and second execution has the same sense, except in links BD and FG. The enhancement produced in the second execution is due to the gradient values of the objective function. In the first execution optimization process has finished because the gradient in all the directions was lower than the adjusted tolerance.

Nevertheless, the second execution the program can optimize better the mechanism. This is because the components of the objective function are normalized with the initial value, and in the second execution the initial value is not the same, and hence, the initial value of the objective function is different too, enabling the gradients are greater than the imposed tolerance.

5 CONCLUSIONS

In this work, a methodology to design robots by considering the mechanical design from the first steps has been developed. The application of the methodology to a robotic arm let to prove the soundness of the work developed in this manuscript. The results conclude that the designed arm, which has a weight around 300 gr. (without considering the actuator system that we estimate will be around 390 gr.) can transport a payload of 1Kg. Better design could be obtained by applying the multi-objective optimization to all the mechanism obtained in the type synthesis process. This is true because the dimensional synthesis has a critical importance in the mechanism kinematic behavior. On other hand, and based on the stress study we can see that the stress level is high due to the optimization performed, by means of the weight reduction.

6 ACKNOWLEDGEMENTS

The financial support of Regional government of Madrid and Carlos III University is gratefully acknowledged through the Project MADBOT 2011/00130/001.

REFERENCES

- [1] Ceccarelli, M. New challenging applications for service robots: problems and experiences. *Proceedings of IARP Workshop on Service Robotics and NanoRobotics* (invited paper, n° 2), Beijing, China, 2009.
- [2] International Federation for Robotics. <http://www.ifr.org/> (last visited 15-03-2011).
- [3] Wu, L., Ceccarelli, M. A numerical simulation for design and operation of an underactuated finger mechanism for LARM hand, *Mechanics Based Design of structures and machines*, 37:86-112, 2009.
- [4] Gao, F., Liu, X.J., Jin, S.L. Optimum design of 3-dof spherical parallel manipulators with respect to the conditioning and stiffness indices, *Mechanism and Machine Theory* 35(9): 1257-1267, 2000.
- [5] Park, J.J., Song, J.B. A nonlinear stiffness safe joint mechanism design for human robot interaction. *Journal of mechanical design* 132(6): 1-8, 2010.
- [6] Carbone, G., Lim, H.O., Takanishi, A., Ceccarelli, M. Optimum Design of a New Humanoid Leg by Using Stiffness Analysis, *12th International Workshop on Robotics in Alpe-Andria-Danube Region RAAD 2003*, Cassino, Italy, 2003.
- [7] Logo J., Mohsen, G., Rad, M.M. Optimal topologies in case of probabilistic loading: The influence of load correlation, *Mechanics Based Design of structures and machines*, 37:327-348, 2009.
- [8] Johnson, R.C. *Mechanical design synthesis: creative design and optimization*, Malabar: Robert E. Krieger Publishing Company, 1978.
- [9] Castejon, C., Carbone, G., García-Prada, J.C., Ceccarelli M. A multi-objective optimization design for a 4R Service Robot, *International Journal on Mechanics and Control* 8(2):3-8, 2007.
- [10] Yan, H.S. *Creative design of mechanical devices*, Singapore: Springer, 1998.
- [11] Bai, S. Optimum design of spherical parallel manipulators for a prescribed workspace, *Mechanism and Machine Theory* 45:200-201, 2010.
- [12] Blanco, D., Castejón, C., Kadhim, S., Moreno, L. Predesign of an Anthropomorphic Lightweight Manipulator, *8th International Conference on Climbing and Walking Robots*, London, U.K, 2005.
- [13] Font-Llagunes, J.M., Kovacs, J. Efficient Dynamic walking: Design strategies to reduce energetic losses of a compass walker at heel strike, *Mechanics based design of structures and machines*, 37: 259-282, 2009.
- [14] Lee, J.H., Nam, Y.J., Park, M.K. Kinematics and optimization of a 2-DOF parallel manipulator with passive constraining leg and linear actuators. *Journal of Mechanical Science and Technology* 24:19-23, 2010.
- [15] Gao, Z., Zhang, D., Ge, Y. Design optimization of spatial six degree-of-freedom parallel manipulator based on artificial intelligence approaches. *Robotics and Computer-Integrated Manufacturing* 26:180-189, 2010.
- [16] Kim, S.M., Kim, W., Yi, B.J. Kinematic analysis and optimal design of a 3T1R type parallel mechanism. *IEEE International conference on robotics and automation*, Kobe, Japan, 2009.
- [17] Yoon, J.W., Ryu, J., Hwang, Y.K. Optimum design of 6-DOF parallel manipulator with translational/rotational workspaces for haptic device application. *Journal of Mechanical Science and Technology* 24(5): 1151-1162, 2010.

- [18] Shigley, J.E., Uicker, J.J. *Theory of Machines and Mechanisms*, Mc Graw-Hill, 2003.
- [19] Merlet, J.P. Optimal design of robots, *Proceedings of Robotics: Science and Systems*, Cambridge, USA, 2005.
- [20] Carbone, G., Ottaviano, E., Ceccarelli, M. An Optimum Design Procedure for Both Serial and Parallel Manipulators, *Proceedings of the Institution of Mechanical Engineers IMechE Part C: Journal of Mechanical Engineering Science* 221(7):829-843, 2007.
- [21] Castejón, C., Carbone, G., García-Prada, J.C., Ceccarelli, M. A Multi-Objective Optimization of a Robotic Arm for Service Tasks. *Journal of Mechanical Engineering Strojivski-Vestnik* 56(5): 316-329, 2007.
- [22] Snyman J.A. *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer Publishing, 2005.
- [23] Dip G., Prahlad V., Kien P.D. Genetic algorithm-based optimal bipedal walking gait synthesis considering tradeoff between stability margin and speed. *Robotica*, 27, 355-365, 2009.
- [24] Coleman, T., Branch, M.A., Grace, A. *Optimization toolbox for use with MATLAB®*. MathWorks Inc, 1999 (available Web http://www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf).

VEHICLE PARAMETERS IDENTIFICATION AND STATES ESTIMATION

Damiano Capra

Mauro Velardocchia

Alessandro Vigliani

Dipartimento di Ingegneria Meccanica e Aerospaziale
Politecnico di Torino

ABSTRACT

This paper presents a method to evaluate vehicle static inertial parameters and dynamic states based on measured wheel forces. By combining a traditional vehicle model with the sensed forces, a simple observer may be devised to compute at first vehicle mass inertial properties; then, using also the other information currently available on passenger cars, the full vehicle state can be estimated. The estimation procedure and algorithms are validated with a nonlinear vehicle model: the results obtained from several simulation examples show that estimates of the inertial parameters generally have small relative errors; moreover also the estimates of the dynamic states appear promising.

Keywords: vehicle parameters; state estimation; vehicle dynamic model; measured hub forces.

1 INTRODUCTION

The design of a global chassis or integrated chassis controller [1, 2] depends on the availability of reliable vehicle state information, which is supplied by the car sensors. In addition, a vehicle state estimator is a crucial algorithm in the controller, with the task of computing additional state and parameter information that can or may not be measured directly. Obviously, limiting the number of sensors increases the performance of the controller at minimum cost. Moreover, also Electronic Control Units (ECU) need state estimators for fail-safe procedures.

The effectiveness of any dynamic stability control system depends on accurate knowledge of the vehicle states, particularly yaw rate and sideslip angle. It is well known that yaw rate can be easily measured with cheap sensors; on the contrary, sideslip angle must be estimated by more sophisticated means. Stability systems currently available on passenger cars typically derive this value from integration of inertial sensors; this method is affected by uncertainty and errors, e.g., due to accumulate sensor errors and unwanted measurements from road grade and bank angle.

An alternative method overcomes some problems by means of Global Positioning System (GPS) measurements [3].

While absolute GPS measurements eliminate the errors from inertial navigation system (INS) integration, INS sensors complement the GPS measurements by providing higher update rate estimates of the vehicle states. However, when GPS signal is lost, as may occur in urban driving environments, integration errors can still accumulate and lead to errors in the estimation process [3].

Accurate knowledge of vehicle states is crucial; consequently, researchers have proposed many different approaches for their estimation. However, most of these approaches are not without their restrictions and limitations; many of the estimators work only with a reduced number of states or a simplified vehicle model, such as the single-track model. Approaches can be found that use this model for the estimation of side slip angle and yaw rate or for lateral acceleration, yaw rate and tyre slip angles, as reported in [4].

Other approaches use more complex vehicle models, but usually have the drawback that the implemented estimator requires the use of a whole set of additional costly sensors. Moreover, all the above approaches show another limitation, i.e., they use fixed vehicle parameters, such as mass, moments of inertia and/or position of the centre of gravity (CoG), which are considered to be known (e.g., the mass of an empty vehicle). Obviously these parameters may vary significantly for each journey: therefore it is evident

Contact author: A. Vigliani¹

¹C.so Duca degli Abruzzi, 24 – 10129 Torino – Italy.

E-mail: mauro.velardocchia@polito.it
alessandro.vigliani@polito.it

that appropriate knowledge of these vehicle parameters is of importance for effective estimation of the various vehicle states. Recognising the need for such knowledge, the estimator described in this paper additionally estimates the parameters to achieve an improvement of the vehicle states. Nam et al. [5] present new real-time methods for the lateral vehicle velocity and roll angle estimation, based on lateral tire forces, obtained from a multi-sensing hub (MSHub) unit.

Tire forces that act during severe manoeuvres are nonlinear with respect to wheel slip and wheel slip angles and cause a wide range of stability and handling properties. The forces also depend on uncontrollable external factors, such as road/tire interface properties, tire pressure and wear, and vehicle loads, that are difficult to sense or anticipate. Because direct sensing of tire forces, slip, slip angles, and external factors is an arduous task, vehicle control systems designed to maintain stability during emergencies typically follow rules based on extensive testing, rather than analytic control laws based on actual forces and accurate state estimates [2].

Integrated or global chassis control aims to enhance driving stability, safety and comfort under all driving conditions through the optimal distribution of lateral, longitudinal and vertical tyre forces acting on the vehicle [2].

A vehicle state estimator as a part of the control system is essential to build a fully integrated control system for braking, chassis and steering using only a limited number of sensors. Currently most of the vehicle state estimators employ Kalman algorithms and use a generic Kalman Filter for estimation based on a linear vehicle model. The Extended Kalman Filter (EKF) and its improved algorithms can be used for non-linear vehicle models [4, 6].

Nowadays passenger cars carry a relevant number of sensors, due to the large number of devices installed, each of them requiring a lot of information. Usually, many sensors provide information for more than one system mainly in order to reduce the costs. To this aim, many sensors are optimized or at least adapted in term of position or characteristic so that they can provide data for different devices but, even with this effort, the number is still high.

The goal of a integrated chassis control system is also to reduce the number of sensors used on the vehicle [1].

Aiming at this target, this paper tries to exploit the best from the Load Sensing Hub Bearing Unit (LS-HBU) [7, 8] or of other sensors, e.g., accelerometers embedded in the tires, which should be capable of measuring the forces exchanged between ground and tires.

The idea at the basis of the paper is to develop an algorithm able to identify some vehicle static parameters, such as mass, moments of inertia and position of the centre of gravity (CoG), as well as some dynamic information, such as yaw rate, during normal driving condition.

This task is accomplished by means of two different algorithms, one devoted to Vehicle Parameters Identification (VPI), the other to Vehicle State Estimator (VSE), both based on the forces exchanged in the hub bearings, which represent the main inputs of the procedure.

The software described in the present paper gives very good results, even if some limits in the estimation cannot be directly overcome: indeed it is impossible to estimate correctly roll and pitch motion without some information about suspensions, since sway-bars, anti-dive and anti-lift have no effect on the forces measured at the wheel hubs. The static part only requires a few geometric parameters, which can be simply measured; instead the information about suspensions is more complex, as it requires some refined measurement.

The tool is divided in two parts: static and dynamic. The first one intervenes when the vehicle is switched on and in the few instants required to start the engine, all the forces are measured and the VPI calculates the position of the center of gravity and the real mass of the car.

With this information, the second part (i.e., the dynamic VSE) can start its task, calculating many dynamic data that can be used by the car ECU to improve the monitoring and control performance.

At the moment, it appears impossible to rely only on the forces measured by hub bearing sensors; in fact some additional data from the car Controller Area Network (CAN) are also necessary. Moreover, since some data do not change during the vehicle lifetime, the estimator can treat them as constant parameters. For those parameters presenting relevant changes during the vehicle life, such as for suspension characteristics, some life-dependent worsening effect estimator could be implemented.

It should be also noticed that the static VPI provides information for the dynamic part, but these are not available until the time required for the estimation has elapsed.

2 STATIC VPI

During its lifetime, any vehicle will be loaded in many different configurations that will directly alter the inertial properties of mass, mass moments of inertia and centre of gravity locations. These changes in load and load distribution can be large when compared with the weight of the empty vehicle. These altered inertial parameters will have a direct impact on the handling, ride, braking and traction performance of the vehicles noted above [9-10].

The aim of Static VPI algorithm is to provide a tool able to estimate the most important parameters (e.g., effective mass and centre of gravity position) that affect the dynamic behaviour of a car. This algorithm is able to identify the vehicle parameters in static conditions or rather when the vehicle is at standstill. It is based on a simple vehicle model and takes the measured hub forces and the wheel steering angle as inputs; moreover it requires some geometrical values (wheelbase, track, reference mass and CoG height) and suspensions parameters (linearized elastic characteristic) to be known.

The results are the car effective mass, the coordinates of the centre of gravity (x_G , y_G , z_G) and the longitudinal (α) and transversal (β) road angles.

Fig. 1 plots the forces acting on the vehicle in static and steady state conditions; moreover, the flow chart relative to

both static and dynamic parameters and states estimators is shown in Fig. 2.

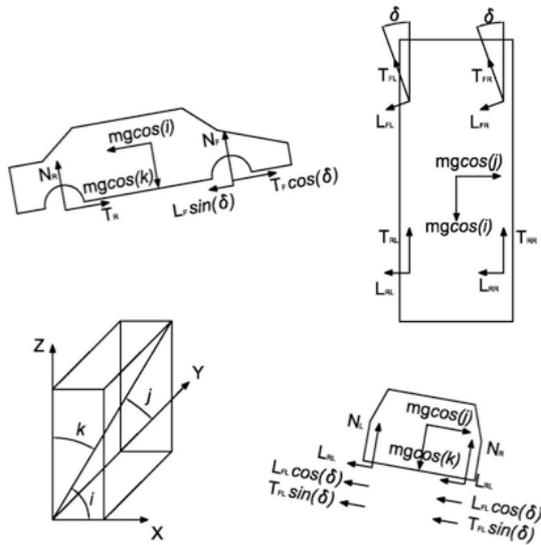


Figure 1 Static forces acting on the vehicle.

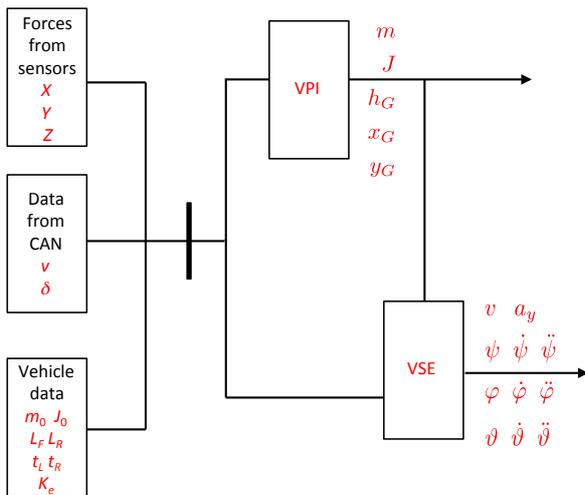


Figure 2 Flow chart of the estimation procedure.

Since the vehicle may have any inclination, the vector defined by the body mass multiplied by the gravity acceleration requires cosine directors i, j and k to be defined in order to write the equations in a general form.

The algorithm is based on the well-known equilibrium equations for a vehicle in stationary conditions (i.e. null velocities and accelerations), as shown in Fig. 1:

$$(X_{FL} + X_{FR}) \cos \delta_w + (X_{RL} + X_{RR}) + (Y_{FL} + Y_{FR}) \sin \delta_w - mg \cos i = 0 \quad (1)$$

$$(Y_{FL} + Y_{FR}) \cos \delta_w + (Y_{RL} + Y_{RR}) + (X_{FL} + X_{FR}) \sin \delta_w - mg \cos j = 0 \quad (2)$$

$$Z_{FL} + Z_{FR} + Z_{RL} + Z_{RR} - mg \cos k = 0 \quad (3)$$

$$L + L_F + L_R = 0 \quad (4)$$

$$t + t_L + t_R = 0 \quad (5)$$

$$h_G - h_{G0} + (m - m_0) / k_e = 0 \quad (6)$$

$$mgL_F \cos k - mgh_G \cos \alpha - (Z_{FL} + Z_{FR})L = 0 \quad (7)$$

$$mgL_R \cos k - mgh_G \cos \beta - (Z_{RL} + Z_{RR})L = 0 \quad (8)$$

$$(Y_{FL} + Y_{FR})L \cos \delta_w + (X_{FL} + X_{FR})L \sin \delta_w - Y_{FR}t \sin \delta_w + X_{RR}t + X_{FR}t \cos \delta_w - mgL_R \cos j - mgt_L \cos i = 0 \quad (9)$$

where k_e is the equivalent vertical stiffness of the vehicle $k_e = \sum k_i$, k_i being the spring stiffness.

The system of equations can be solved by means of iterative techniques: Tab.1 reports some results obtained by the static VPI. Results from simulations appear excellent: even in presence of errors or disturbances on the measured wheel forces of about 10%, the estimated values of the vehicle parameters show errors smaller than 5%, while the estimate of the road side and longitudinal slope is smaller than 12%.

Table I - Real vs. estimated parameters

Data	Real value	Estimated value	Variation [%]
m [kg]	1582	1582.1	0.01
h_G [m]	0.49	0.485	-1.02
L_F [m]	1	0.98	-2.00
L_R [m]	1.7	1.72	1.18
t_L [m]	0.785	0.785	0.00
t_R [m]	0.785	0.785	0.00
α [deg]	10	11	10.00
β [deg]	1	0.95	5.00

2.1 ESTIMATION OF THE MASS MOMENTS OF INERTIA

The static VPI has been developed to estimate also the mass moments of inertia of the vehicle. The values obtained are rough estimation of the real ones, because there no dynamic effect is considered; nevertheless, the estimate is sufficient to provide some data to improve the performances of the dynamic VSE. Obviously, results could improved if dynamic conditions are taken into account, as suggested in [10].

The vehicle information necessary to the static VPI for estimating the mass moments of inertia, are the mass m and the centre of gravity position G . In particular it is possible to store into the control unit the nominal value of the mass m_0 and CoG position without driver and baggage, as measured at the end of production line.

The value of the moment of inertia consequent to the presence of an additional mass Dm that shifts the vehicle CoG into a new position C can be estimated by means of the following formula:

$$J_C = J_0 + md^2(2m - m_0) / (m - m_0) \quad (10)$$

where the distance $d = CG$ can be estimated by the VPI procedure.

As an example, let consider a vehicle with a mass moment of inertia $J_0 = 2210 \text{ kgm}^2$: this value is the reference moment of inertia inserted in the algorithm. An additional mass (60 cm x 120 cm x 50 cm) of 150 kg is placed at a distance $d = 1 \text{ m}$ from the CoG, thus changing the vehicle mass moment of inertia J_C about the vertical z axis. Tab. 2 shows the result of the estimate: the difference between estimated and actual values is very small.

Table II - Real vs. estimated mass moment of inertia

Data	Real value	Estimated value	Variation [%]
$J [\text{kgm}^2]$	2382	2404	0.94

3 DYNAMIC VSE

The goal of the dynamic VSE is to provide data concerning the vehicle behaviour to the ECUs on the vehicle; the proposed algorithm is based on the classical equations of motion (12), which are reported in Appendix B. They can be written in matrix form as:

$$[A]\{q\} = \{F_e\} \quad (11)$$

where $\{q\}$ is the unknown vector of angles α and β , lateral acceleration a_y , and pitch, roll and yaw angular accelerations $(\ddot{\varphi}, \ddot{\theta}, \ddot{\psi})$.

In particular, numerical tests show that yaw acceleration and velocity are very well approximate: as visible in Fig. 3 and 4, the curves obtained from the estimate and from a 14 dof vehicle model are almost coincident. Obviously, since yaw rate is the integration of the acceleration, so some drifting tendency appears after 2.5 s

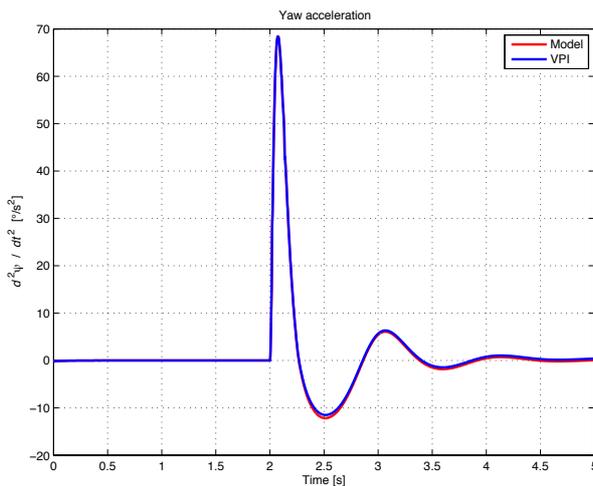


Figure 3 Yaw acceleration: comparison between modelled and estimated value.

As it can be observed from Fig. 5, also the lateral acceleration is well estimated: due to integration, accuracy is lost after 2.8 s; nevertheless, the difference between simulated and estimated values is sufficiently small so that it represents a reliable input for any stability control system.

The simulated results are obtained using a 14 degrees of freedom nonlinear vehicle model.

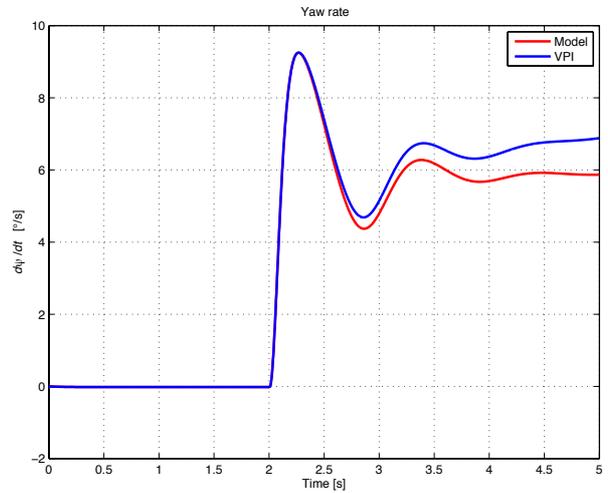


Figure 4 Yaw rate: comparison between modelled and estimated value.

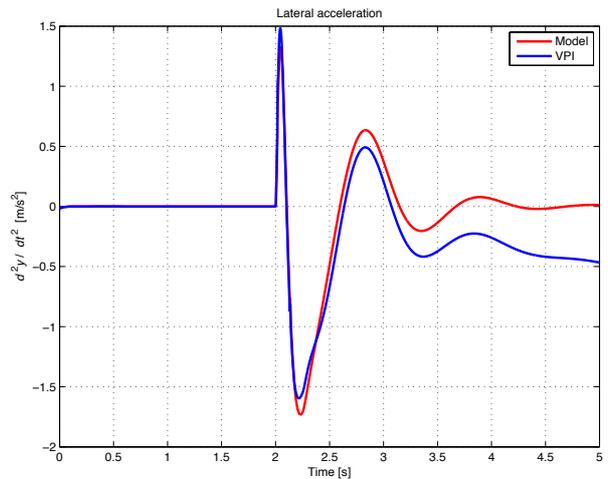


Figure 5 Lateral acceleration: comparison between modelled and estimated value .

2.1 PITCH AND ROLL ESTIMATION

Equations (12) give good estimates of all the variables describing the vehicle behaviour except for pitch and roll motion, due to the effects of anti-dive/anti-lift and sway bars. In fact, the geometry of the suspension affects the pitch angle of the chassis though it has no influence on the forces measured at the hub bearings; hence, the estimation from LS-HBU measurements is not able to approximate the real value. Similarly, the sway bars generate an anti-roll torque acting directly on the sprung mass, without affecting the forces at the hubs.

An obvious solution is to implement a model able to simulate the anti-dive/anti- lift and sway bar effects, in contrast with the VPI aim: in fact, the tool should be able to

work on any vehicle, without devoting energy and time in the model of the previous effects, which strongly depend on the vehicle model.

Consequently, in order to obtain better estimates without losing generality, two different approaches have been tested. The first consists of modelling the previous effect by means of a torsional spring and damper: it is a simple method, but it requires noticeable calibration effort to obtain good results (note that it is not able to approximate perfectly the dynamic behaviour).

The second method relies on an adaptive filter, able to correct itself in order to better approximate the solution.

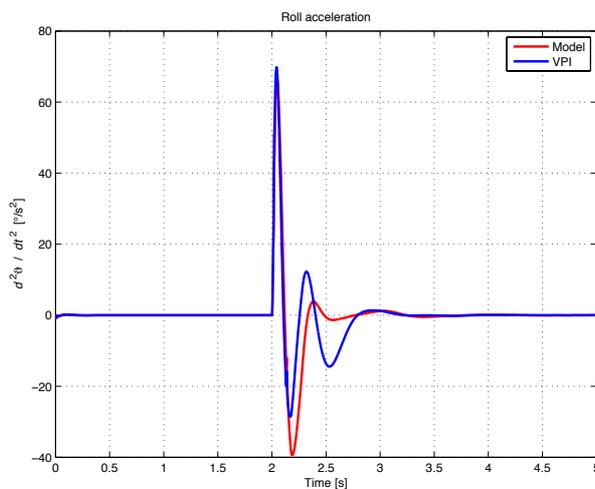


Figure 6 Roll acceleration: comparison between modelled and estimated value.

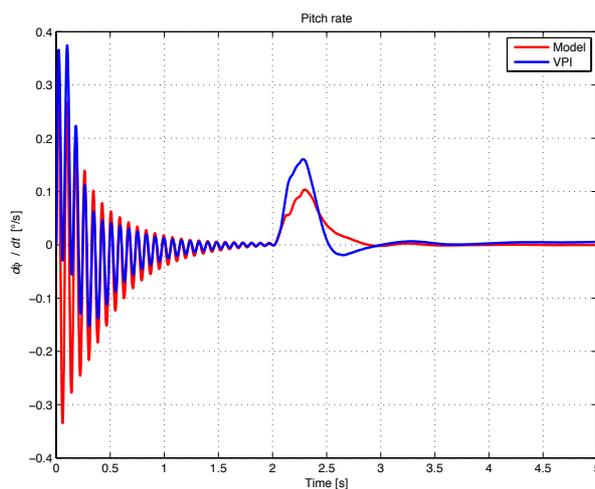


Figure 7 Pitch rate: comparison between modelled and estimated value.

The most famous is Kalman's filter. The equations that should be implemented into the filter are those related to pitch and roll. The procedure is not simple, due to the fact that the desired variables cannot be explicitly obtained from the equations of motion. One solution is to operate a

linearization around a certain point, which can change every time step, so as to have a linear system. To this aim it is necessary to evaluate the Jacobian of the system: in this paper two plots (Fig. 6-7) showing the numerical results are presented.

4 CONCLUSIONS

This article has shown a possible application of Load Sensing Hub Bearing: they can be used to provide the force inputs for realising a model-based vehicle state and parameter estimator. It must be noted that the proposed approach requires some parameters (curb weight, linearised suspension characteristics, etc.) to be known; moreover some additional measures are still needed from other sensors in order to obtain satisfactory and reliable performances.

One advantage of this technique is that with an efficient estimator and an improved accuracy of the parameter estimates, the model uncertainties within the state estimator decrease, thus allowing an enhancement of the quality of the state estimation. A further benefit is the possibility to switch off the parameter estimator, once the parameter estimates converge to satisfactory values or when the signal/noise ratio indicates disturbed input and output signals. In this case, constant parameter estimates are maintained, thus reducing the uncertainties for the state estimator.

The estimator is implemented in the Matlab/Simulink software environment and results are compared with those obtained with a 14 degrees of freedom nonlinear vehicle model.

Another advantage lies in the principle of replacing the existing individual state estimators, required by the various stability control systems within a vehicle, by one comprehensive single state and parameter estimator.

Future work will investigate how to apply this information to extend the ability of the observer to predict vehicle motion throughout the entire range of handling behaviour. Moreover it will evaluate the errors due to the sensor delays and uncertainties, comparing also other sensors that can be used for the measurement of tire forces.

Further work is planned to extend the estimator by increasing the number of states and parameters, whilst at the same time improving stability and computational efficiency. Of particular interest is the inclusion of the vehicle's environment, such as the road friction, which could be treated as an additional state to be estimated.

REFERENCES

- [1] Yu F., Li D.F. and Crolla D.A., *Integrated Vehicle Dynamics Control – State of the art review*. (VPPC), Harbin, China, 2008.
- [2] Ray L.R., *Nonlinear State and Tire Force Estimation for Advanced Vehicle Control*. *IEEE Trans. Control System Technologies*, Vol. 3, No. 1, pp. 127–124, 1995.

- [3] Ryu J., Rossetter E., and Gerdes J.C., Vehicle sideslip and roll parameter estimation using GPS. *Proc. Int. Symposium on Advanced Vehicle Control (AVEC)*, Hiroshima, Japan, 2002.
- [4] Wenzel T.A., Burnham K.J., Blundell M.V. and Williams R.A., Dual extended Kalman filter for vehicle state and parameter estimation. *Vehicle System Dynamics*, Vol. 44, No. 2, pp. 153–171, 2006.
- [5] Nam K., Oh S., Fujimoto H. and Hori Y., Vehicle State Estimation For Advanced Vehicle Motion Control Using Novel Lateral Tire Force Sensors. *American Control Conference*, San Francisco, CA, USA, 2011.
- [6] Venhovens P.J.T. and Naab K., Vehicle dynamics estimation using Kalman filters. *Vehicle System Dynamics*, Vol. 32, No. 2–3, pp. 171–184, 1999.
- [7] Zuurbier J. and van Leeuwen B.G., Vehicle dynamics control based on force-sensing wheel bearings. *Vehicle Dynamics Expo 2007*, Messe Stuttgart, Germany, 2007.
- [8] Morgando A., Ondrak V., van Leeuwen B.G., Velardocchia M. and Vigliani, An alternative approach to automotive ESC based on measured wheel forces. *Vehicle System Dynamics*, Vol. 49, No. 12, pp. 1855–1871, 2011.
- [9] Fathy H.K., Kang D. and Stein J.L., Online Vehicle Mass Estimation Using Recursive Least Squares and Supervisory Data Extraction. *American Control Conference*, Seattle, Washington, USA, 2008.
- [10] Rozyn M. and Zhang N., A method for estimation of vehicle inertial parameters, *Vehicle System Dynamics*. Vol. 48, No. 5, pp. 547–565, 2010.

APPENDIX A

Nomenclature

- α : road longitudinal inclination angle
 β : road lateral inclination angle
 φ : pitch angle
 θ : roll angle
 ψ : yaw angle
 h_G : center of gravity effective height
 h_{G0} : center of gravity reference height
 k_e : equivalent spring stiffness
 m : actual mass
 m_0 : curb mass
 t : vehicle track
 t_L : left halftrack
 t_R : right halftrack
 L : wheelbase
 L_F : front wheelbase
 L_R : rear wheelbase
 J : mass moment of inertia
 X : longitudinal tire force
 Y : side tire force
 Z : vertical tire force

subscripts:

- d : drag
 s : sprung
 w : wheel
 FL : front left
 FR : front right
 RL : rear left
 RR : rear right

APPENDIX B

Nonlinear equations of motion of the six degrees of freedom sprung mass model

$$\begin{aligned}
 m_s \ddot{x} + m_s g \cos \beta \sin \alpha + m_s h_G \cos \vartheta \cos \varphi \ddot{\varphi} + m_s h_G \sin \vartheta \ddot{\psi} &= (X_{FL} + X_{FR}) \cos \delta_w - (Y_{FL} + Y_{FR}) \sin \delta_w + X_{RL} + X_{RR} + F_d + m_s \dot{y} \dot{\psi} \\
 m_s \ddot{y} - m_s h_G \cos \vartheta \ddot{\varphi} + m_s h_G \cos \vartheta \sin \varphi \ddot{\psi} &= (X_{FL} + X_{FR}) \sin \delta_w + (Y_{FL} + Y_{FR}) \cos \delta_w + Y_{RL} + Y_{RR} - m_s \dot{x} \dot{\psi} \\
 m_s g \cos \alpha \cos \beta - m_s h_G \sin \vartheta \cos \varphi \ddot{\vartheta} - m_s h_G \cos \vartheta \sin \varphi \ddot{\varphi} &= Z_{FL} + Z_{FR} + Z_{RL} + Z_{RR} \\
 (m_s h_G^2 + J_{xx}) \ddot{\vartheta} - m_s h_G \cos \vartheta \ddot{y} &= (Z_{FL} + Z_{FR}) t_L - (Z_{RL} + Z_{RR}) t_R + m_s h_G \dot{x} \dot{\psi} \\
 m_s g h_G \sin \alpha \cos \beta + (m_s h_G^2 \cos^2 \vartheta + J_{yy} \cos^2 \vartheta) \ddot{\varphi} &= (Z_{FL} + Z_{FR}) L_F - (Z_{RL} + Z_{RR}) L_R - m_s h_G \ddot{x} \\
 m_s h_G \cos \vartheta \sin \varphi \ddot{y} + J_{zz} \cos^2 \vartheta \cos^2 \varphi \ddot{\psi} &= (-X_{FL} \cos \delta_w + Y_{FL} \sin \delta_w) t_L + (X_{FR} \cos \delta_w - Y_{FR} \sin \delta_w) t_R + (X_{FL} + X_{FR}) L_F \sin \delta_w + \\
 &+ (Y_{FL} + Y_{FR}) L_F \cos \delta_w - X_{RL} t_L + X_{RR} t_R - (Y_{RL} + Y_{RR}) L_R
 \end{aligned} \tag{12}$$

APPLICATION OF A NEW TAILORING OPTIMIZATION TECHNIQUE TO LAMINATED AND SANDWICH PLATES AND TO SANDWICH SPHERICAL PANELS

U. Icardi F. Sola

Department of Mechanical and Aerospace Engineering - Politecnico di Torino
Corso Duca degli Abruzzi 24 – 10129 Torino, Italy

ABSTRACT

A tailoring optimization technique recently developed for improving structural response and energy absorption of composites is applied to laminated and sandwich plates and to sandwich spherical panels. The in-plane variation of the stiffness properties of plies and the through-the-thickness variation of the core properties are determined solving the Euler-Lagrange equations of an extremal problem in which the strain energy due to out-of-plane strains and stresses is minimized, while that due to their in-plane counterparts is maximized. The zig-zag model with hierarchic representation of displacements recently developed by the authors is used as structural model. It accurately describes the strain energy and the interlaminar stresses directly from the constitutive equations. It *a priori* fulfils the displacement and stress contact conditions at the interfaces and considers a hierarchic representation that adapts itself to the variation of solutions. Refinement of the solution does not imply an increase of the number of functional degrees of freedom, which are the traditional mid-plane displacements and the shear rotations. As shown by the numerical applications, the critical interlaminar stress concentrations at the interfaces are sensibly reduced without any bending stiffness loss and the strength to debonding of faces from the core is improved.

Keywords: tailoring optimisation; laminated and sandwich structures; zig-zag model

1 INTRODUCTION

Fibre reinforced and sandwich composites with laminated faces are advantageous respect conventional metallic structures, thanks to better specific strength and stiffness, good behaviour under fatigue, great protection against fragments' projection, survivability, containment of explosions, impact resistance, and vibration suppression. In addition these materials present many other favourable properties, here not cited, but above all they also offer the remarkable advantage of being designed to fulfil specific requirements, thank to the tailoring technique.

These materials, unfortunately, absorb a large amount of the incoming energy through local failures that usually cause detrimental effects, such as an embedded delamination.

In order to fully exploit the potential advantages of these materials it is mandatory an accurate assessment of the local damage mechanisms and post failure behaviour.

The reader is referred to the review papers by Rowlands [1], Tennyson and Wharam [2]; Nahas [3]; Bolotin [4]; Echaabi and Trochu [5]; Paris [6]; Icardi et al. [7] for a comprehensive discussion of the local failure mechanisms and residual properties of these materials.

One of the advantage of the composites materials is the possibility of be tailored; in particular the structural performance of laminated structures may be optimised by finding the appropriate ply angles and stacking sequence that maximize design requirements. As examples of ply angles optimization studies and of sandwich optimization, the papers by Chaperon et al. [8], Jones et al. [9], Keller [10], Kaye and Heller [11] and Sivakumar et al. [12] are cited. Customary objective function and constraints are evaluated using finite element or closed form analytic approaches, and in order to efficiently solve the optimal lamination problem, gradient based search techniques or genetic algorithms are employed (see, e.g. Uys et al. [13] and Krishnapillai and Jones [14], respectively).

Contact author: Icardi U.¹

Corso Duca degli Abruzzi 24, 10129 Torino, Italy
¹ugo.icardi@polito.it

The structural response is generally computed using smeared laminate models, since the computational effort, required by more realistic models, is unaffordable for the optimisation process. So with this approach, the layerwise effects that adversely affect strength, stiffness, failure behaviour and service life of composites may not be accurately predicted.

In addition many optimization process are carried out considering the fibre orientation angle constant throughout the plies and assuming the core properties uniform across the thickness; in this way an excellent opportunity for improving performance and damage tolerance of sandwiches and for enlarging their life in service may not be fully exploited. In fact with this approach the interlaminar stress concentration due to the different mechanical properties of faces and core may not be recovered. Therefore, debonding can lead to stiffness loss, bending and buckling failure in service.

Variable-stiffness composites in which the fibres follow curvilinear paths have been considered by Hyer and Charette [15] and Hyer and Lee [51] since the last two decades. The aim is to vary at each point the properties, to enable the fulfilment of specific design requirements. This tailoring option offers many advantages over straight-fibre laminates such as the improvement of buckling and first-ply failure loads (Jegley et al. [16] and Lopes et al. [17]), the maximisation of stiffness (Pedersen [18] and Setoodeh et al. [19]) and an increase of the fundamental frequency (Diaconu et al. [20], Narita and Hodgkinson [21] and Abdalla et al. [22]). For these reasons Ghiasi et al. [23] and Javidrad and Nouri [24] have recently retaken this technique. As Jung [25] and Lakes [26] demonstrated, it is possible to obtain the desired structural properties without any stiffness loss, by using structural hierarchy guided combinations of materials with different stiffness/dissipation properties.

As far as manufacturing technologies are concerned, the advent of variable stiffness composites is allowed thanks to technologies such as automated fibre-placement (Barth [27], Enders and Hopkins [28], Martin et al. [29] and Evans [30]) and functionally graded materials (FGM) (see, e.g., Fuchiyama and Noda [31], Suresh and Mortensen [32] and Mahfuz et al. [33]). In particular the great potential advantages of variable-stiffness materials have been recently proven owing to the studies on FGM by Sankar [34], Sankar and Tzeng [35], Venkataraman and Sankar [36], Apetre et al. [37], [38], [39], Zhu and Sankar [40], Reid and Paskaramoorthy [41].

Despite the great advantages of variable stiffness composites, the possibility of varying the strength and stiffness properties at each point determines a significantly enlarged number of design variables. For this reason the traditional optimization techniques based on gradient search or genetic algorithms may become computationally too expensive with variable-stiffness composites, making unpractical these materials. To overcome this problem a tailoring optimisation technique based on variable-stiffness

concepts (OPTI) has been developed progressively refined in Refs. [42], [43], [44] and [45] for plane structures and then extended to shell structures in Ref. [52].

This technique solves the Euler–Lagrange equations that follow by imposing the first variation of the strain energy to zero under spatial variation of the stiffness properties, in this way the optimal fibre angle variation of the face is determined. With this approach, the optimisation problem of variable-stiffness composites is reduced to a simple problem of finding the appropriate stacking sequence like with straight-fibre composites, which can be efficiently solved by the classical optimisation techniques. On the contrary the classical optimization techniques may overwhelm the computational capacity, due to the huge number of design variables involved by the variation of the properties at any point. Since laminated and sandwich composites have a much smaller strength and stiffness in the thickness direction than in the in-plane direction, the aims of OPTI are minimising the energy stored by the modes involving large out-of-plane strains and stresses and maximising that of modes with dominant membrane strains and stresses. The merits of this technique have been demonstrated in many works; in Ref. [42], applications were presented to laminated plates subject to low velocity, low energy impacts, which proven that beneficial effects may be obtained on the strength at the onset of delamination incorporating layers with spatially variable optimised properties. Ref. [43] considered sandwich flat panels with through-the-thickness variable distributions of the core properties and variable in-plane distributions of the face plies properties under blast pulse loading, to show that wanted dynamic response properties can be achieved recovering the critical interlaminar stresses. Ref. [44] carried out the optimisation of sandwich flat panels using a finite element scheme based on strain energy updating concepts, while Ref. [45] showed the improved “tuning” capability of double-core sandwich flat panels over single core sandwiches. These studies were achieved using the structural model of Ref. [46], which is a 3D zig-zag plate model with a third-order piecewise variation of the in-plane displacements and a fourth-order piecewise variation of the transverse displacement across the thickness which *a priori* fulfils the interlaminar stress contact conditions at the interfaces. Ref. [52] is an extension of OPTI to shells, where for the first time applied an adaptive model to optimization problems.

This paper presents an application of OPTI to laminated plates and spherical panels, not still considered. In particular the goal is finding the spatial variation of stiffness properties of the layers that minimise the energy due to interlaminar stresses and maximises the bending stiffness, since composites have weak out-of-plane properties, as mentioned above. To this purpose an accurate description of the strain energy is mandatory, according this study is carried out by employing the multi-layered 3D zig-zag model of Ref. [47], which is here extended to shell structures.

This model is characterised by a hierarchic representation of the displacements across the thickness. In order to have the minimal number of primary variables, the model assumes as functional degrees of freedom (d.o.f.) the classical mid-plane displacements and shear rotations.

The main advantage of this model is that it can provide accurate stress predictions directly from constitutive equations, without using post-processing technique, such as the integration of local differential equations, which can fail in capturing accurate interlaminar stresses, as shown by Cho et al. [48], Li and Liu [49] and Zhen and Wanji [50].

The paper is structured as follows. First of all a brief discussion over the features of the structural model and the basic steps towards its development are presented. Next the process for deriving the Euler-Lagrange equations which represent the stationary conditions under variation of the stiffness properties is reviewed, and finally numerical applications to laminated plates and spherical panels are presented.

2 THE STRUCTURAL MODEL

Summing up, two categories of multi-layered structural models for analysis of laminated and sandwich composites can be detected, those with a variable number of functional d.o.f. or those with a fixed number of functional d.o.f. The former can be refined across the thickness and use a different representation in different computational layers, so in order to improve accuracy a large number of computational layer have to be employed. For this reason their application requires a large computational effort, thus these models should be limited to local analyses for not overwhelming the computational capacity. On the contrary the structural models with a fixed number of functional d.o.f. cannot be refined across the thickness, so they have a limited validity for stress analysis and employing post-processing techniques have to be employed for improving their accuracy.

The middle surface Ω is assumed as the shell reference surface and the curvilinear tri-orthogonal system constituted by the lines α , β of principal curvature and the coordinate across the thickness ζ as the reference system.

The symbols u_α , u_β , u_ζ represent respectively the elastic displacements in the direction of α , β and ζ , while the radii are denoted as $R_\alpha(\alpha, \beta)$ and $R_\beta(\alpha, \beta)$.

The Lamé coefficients on Ω are indicated as $A_\alpha(\alpha, \beta)$, $A_\beta(\alpha, \beta)$; their counterparts at a distance ζ are indicated as $H_\alpha^{-1}(\alpha, \beta, \zeta)$, $H_\beta^{-1}(\alpha, \beta, \zeta)$. The position of the upper⁺ and lower⁻ surfaces of the k^{th} layer are indicated as ${}^{(k)}\zeta^+$ and ${}^{(k)}\zeta^-$; the quantities that belong to a generic layer k are denoted with the suffix ${}^{(k)}$.

The shell is made by N orthotropic layers, perfectly bonded together, with material properties entirely different each other and with their principal material directions arbitrarily oriented with respect to α . This representation is suited also for sandwich shells, which are treated as multilayered shells made of an arbitrary number of thin layers constituting the faces and of a thick intermediate layer constituting the honeycomb core, whose properties are computed from the cellular properties (see, e.g., Gibson and Ashby [53]).

The differentiation with respect to the spatial coordinates is denoted by the symbols $()_{,\alpha}$, $()_{,\beta}$.

As in the previous papers by the authors, the expressions of the strain field is derived from the displacements field using linear strain-displacement relations, since the displacements are supposed to remain always small:

$$\varepsilon_{\alpha\alpha} = \frac{u_{\alpha,\alpha}}{H_\alpha} + \frac{H_{\alpha,\beta}}{H_\alpha H_\beta} u_\beta + \frac{H_{\alpha,\zeta}}{H_\alpha H_\zeta} u_\zeta \quad (1)$$

$$\varepsilon_{\beta\beta} = \frac{u_{\beta,\beta}}{H_\beta} + \frac{H_{\beta,\alpha}}{H_\alpha H_\beta} u_\alpha + \frac{H_{\beta,\zeta}}{H_\beta H_\zeta} u_\zeta \quad (2)$$

$$\varepsilon_{\alpha\beta} = \frac{H_\alpha}{H_\beta} \left(\frac{u_\alpha}{H_\alpha} \right)_{,\beta} + \frac{H_\beta}{H_\alpha} \left(\frac{u_\beta}{H_\beta} \right)_{,\alpha} \quad (3)$$

$$\varepsilon_{\alpha\zeta} = \frac{H_\alpha}{H_\zeta} \left(\frac{u_\alpha}{H_\alpha} \right)_{,\zeta} + \frac{H_\zeta}{H_\alpha} \left(\frac{u_\zeta}{H_\zeta} \right)_{,\alpha} \quad (4)$$

$$\varepsilon_{\beta\zeta} = \frac{H_\beta}{H_\zeta} \left(\frac{u_\beta}{H_\beta} \right)_{,\zeta} + \frac{H_\zeta}{H_\beta} \left(\frac{u_\zeta}{H_\zeta} \right)_{,\beta} \quad (5)$$

$$\varepsilon_{\zeta\zeta} = \frac{1}{H_\zeta} u_{\zeta,\zeta} \quad (6)$$

with $1/H_\alpha$, $1/H_\beta$ that are approximated up to the second-order, while $H_\zeta = 1$:

$$\frac{1}{H_\alpha} = \frac{1}{A_\alpha} \left(1 - \frac{\zeta}{R_\alpha} + \left(\frac{\zeta}{R_\alpha} \right)^2 \right) \quad (7)$$

$$\frac{1}{H_\beta} = \frac{1}{A_\beta} \left(1 - \frac{\zeta}{R_\beta} + \left(\frac{\zeta}{R_\beta} \right)^2 \right) \quad (8)$$

then the stresses are written using anisotropic 3D stress-strain relations:

$$\sigma_{\mu\nu}^{(k)} = \overline{Q}_{\mu\nu\gamma\delta}^{(k)} \varepsilon_{\gamma\delta}^{(k)} \quad \sigma_{\mu\zeta}^{(k)} = \overline{Q}_{\mu\zeta\nu\zeta}^{(k)} \varepsilon_{\nu\zeta}^{(k)} \quad (9)$$

with μ, ν, γ, δ that in turn represent α, β (repeated indices imply the Einsteinian summation convention).

The following piecewise representation of the displacements is postulated across the thickness:

$$u_\alpha(\alpha, \beta, \zeta) = u_\alpha^0(\alpha, \beta) \left(1 + \frac{\zeta}{R_\alpha} \right) + \zeta(\gamma_\alpha^0(\alpha, \beta) - u_{\zeta, \alpha}^0) + C_\alpha^i(\alpha, \beta)\zeta^2 + D_\alpha^i(\alpha, \beta)\zeta^3 + (O\zeta^4 \dots) + \sum_{k=1}^{n_i} \Phi_\alpha^k(\alpha, \beta)(\zeta - \zeta_k)H_k + \sum_{k=1}^{n_i} C_u^k(\alpha, \beta)H_k \quad (10)$$

$$u_\beta(\alpha, \beta, \zeta) = u_\beta^0(\alpha, \beta) \left(1 + \frac{\zeta}{R_\beta} \right) + \zeta(\gamma_\beta^0(\alpha, \beta) - u_{\zeta, \beta}^0) + C_\beta^i(\alpha, \beta)\zeta^2 + D_\beta^i(\alpha, \beta)\zeta^3 + (O\zeta^4 \dots) + \sum_{k=1}^{n_i} \Phi_\beta^k(\alpha, \beta)(\zeta - \zeta_k)H_k + \sum_{k=1}^{n_i} C_v^k(\alpha, \beta)H_k \quad (11)$$

$$u_\zeta(\alpha, \beta, \zeta) = u_\zeta^0(\alpha, \beta) + b^i(\alpha, \beta)\zeta + c^i(\alpha, \beta)\zeta^2 + d^i(\alpha, \beta)\zeta^3 + e^i(\alpha, \beta)\zeta^4 + (O\zeta^5 \dots) + \sum_{k=1}^{n_i} \Psi^k(\alpha, \beta)(\zeta - \zeta_k)H_k + \sum_{k=1}^{n_i} \Omega^k(\alpha, \beta)(\zeta - \zeta_k)^2 H_k + \sum_{k=1}^{n_i} C_w^k(\alpha, \beta)H_k \quad (12)$$

The superscripts *i* mean that these terms are valid only in a single layer of the laminate.

The representation can be refined across the thickness, without increasing the number of primary variables, by suitably choosing the contributions ($O\zeta^4 \dots$) and ($O\zeta^5 \dots$), as outlined hereafter.

The terms in the summations are the zig-zag contributions whose purpose is to make discontinuous the derivatives of displacements at the layer interfaces in order to fulfil *a priori* the continuity of displacements, of transverse shear, of normal stresses and of the transverse normal stress gradient, as prescribed by the elasticity theory for keeping equilibrium. The Heaviside unit step function H_k makes active the contribution of Φ_α^k , Φ_β^k , Ψ^k , Ω^k , C_u^k , C_v^k , C_w^k since their pertinent interface, i.e. $H_k=1$ for $\zeta \geq \zeta_k$ and 0 for $\zeta < \zeta_k$.

The functional d.o.f. of the present plate model are the three displacements u_α^0 , u_β^0 , u_ζ^0 and the two shear rotations γ_α^0 , γ_β^0 of the points on the reference middle surface of the plate. The coefficients C_α^i , C_β^i , D_β^i , D_α^i , as well as the other high order coefficients are determined by enforcing equilibrium conditions at discrete points across

the thickness and also by enforcing the stress-free boundary conditions for the transverse shear stresses at the upper $|^u$ and lower $|_l$ free surfaces:

$$\sigma_{\alpha\zeta} |^u = 0 \quad \sigma_{\alpha\zeta} |_l = 0 \quad (13)$$

$$\sigma_{\beta\zeta} |^u = 0 \quad \sigma_{\beta\zeta} |_l = 0 \quad (14)$$

The still unknown coefficients b^i to e^i are determined by enforcing equilibrium conditions at discrete points across the thickness as well as by enforcing the fulfilment of the boundary conditions on the transverse normal stress and its gradient at the upper and lower bounding surfaces:

$$\sigma_\zeta |^u = p^0 |^u \quad \sigma_\zeta |_l = p^0 |_l \quad (15)$$

$$\sigma_{\zeta, \zeta} |^u = 0 \quad \sigma_{\zeta, \zeta} |_l = 0 \quad (16)$$

(p^0 represents the transverse distributed loading).

The expressions of the continuity functions Φ_α^k , Φ_β^k , Ψ^k , Ω^k , C_u^k , C_v^k , C_w^k in terms of the d.o.f. and of their derivatives are determined in a straightforward way by enforcing the continuity of the displacement, of the interlaminar transverse shear and normal stresses and of the transverse normal stress gradient at the layer interfaces:

$$\sigma_{\alpha\zeta}(\zeta = {}^{(k)}\zeta^+) = \sigma_{\alpha\zeta}(\zeta = {}^{(k+1)}\zeta^-) \quad (17)$$

$$\sigma_{\beta\zeta}(\zeta = {}^{(k)}\zeta^+) = \sigma_{\beta\zeta}(\zeta = {}^{(k+1)}\zeta^-) \quad (18)$$

$$\sigma_\zeta(\zeta = {}^{(k)}\zeta^+) = \sigma_\zeta(\zeta = {}^{(k+1)}\zeta^-) \quad (19)$$

$$\sigma_{\zeta, \zeta}(\zeta = {}^{(k)}\zeta^+) = \sigma_{\zeta, \zeta}(\zeta = {}^{(k+1)}\zeta^-) \quad (20)$$

$$u_\alpha(\zeta = {}^{(k)}\zeta^+) = u_\alpha(\zeta = {}^{(k+1)}\zeta^-) \quad (21)$$

$$u_\beta(\zeta = {}^{(k)}\zeta^+) = u_\beta(\zeta = {}^{(k+1)}\zeta^-) \quad (22)$$

$$u_\zeta(\zeta = {}^{(k)}\zeta^+) = u_\zeta(\zeta = {}^{(k+1)}\zeta^-) \quad (23)$$

3 OPTIMIZATION OF MATERIAL PROPERTIES

The strain energy optimisation process has the aim of finding a proper distribution of the stiffness properties that minimises the energy absorbed through unwanted modes (e.g., modes involving interlaminar strengths) and maximises that absorbed by desired modes (e.g., modes involving membrane strengths). This distribution is determined by solving the Euler-Lagrange equations, obtained by making extremal the in-plane, bending and out-of-plane shear contributions to strain energy under spatial variation of the stiffness properties.

In this way the amount of the energy stored by specific modes can be minimised, or maximised, as desired, so this technique works as an energy “*tuning*”.

The strain energy optimisation process here presented follows the procedure already described in Ref. [52] but the model is different. In order to find the stationary condition for the strain energy contributions, as the first step the strain energy of the shell model is written. The procedure starts by imposing zero its first variation under variation of the functional d.o.f.; note that this constraint associated to the variation of the external work represents the equilibrium condition. Enforcing the vanishing of the first variation of the strain energy, with the stiffness properties as primary variables, a set of governing equations is obtained, the so called Euler–Lagrange equations that represent the relations enabling the stiffness properties to make the strain energy contributions extremal. The closed form or the numerical solution of these equations gives the optimal distribution of the stiffness properties. The strain energy in bending is the contribution by the terms that multiply u_ζ^0 , while the terms that multiply $\gamma_\alpha^0, \gamma_\beta^0$ represent the strain energy of transverse shears. The tailoring optimization consists in finding stiffness distributions that simultaneously solve these equations. In this procedure the in-plane contributions multiplying u_α^0, u_β^0 are disregarded since laminated and sandwich composites have a larger strength and stiffness in the in-plane direction than in the thickness direction. However the in-plane stresses will be assessed, as they may cause the failure. Since the tailoring of the face plies and of the core require different treatment, the tailoring optimisation process should be split into two procedure; one, regarding the face plies, which is carried out in α, β , and one, regarding the core, which is carried out in ζ . Considering the face plies, the stationary condition for the bending energy is:

$$[MR] \begin{Bmatrix} \delta u_\alpha^0 \\ \delta u_\beta^0 \\ \delta u_\zeta^0 \\ \delta \gamma_\alpha^0 \\ \delta \gamma_\beta^0 \end{Bmatrix} = 0 \tag{24}$$

Where [MR]:

$$\begin{bmatrix} [MR^\alpha] & 0 & 0 & 0 & 0 \\ 0 & [MR^\beta] & 0 & 0 & 0 \\ 0 & 0 & [MR^\zeta] & 0 & 0 \\ 0 & 0 & 0 & [MR^{\gamma^\alpha}] & 0 \\ 0 & 0 & 0 & 0 & [MR^{\gamma^\beta}] \end{bmatrix} \tag{25}$$

Notice that $[MR^\alpha], \dots, [MR^{\gamma^\beta}]$ have different dimensions.

Eq. (24) is obtained integrating by parts the virtual variation of the strain energy under variation of the functional d.o.f. and collecting the contributions that multiply each displacement d.o.f.

The stationary condition for the transverse shear energy in the plane (α, ζ) is obtained with the same procedure as:

$$[SR] \begin{Bmatrix} \delta u_\alpha^0 \\ \delta u_\beta^0 \\ \delta u_\zeta^0 \\ \delta \gamma_\alpha^0 \\ \delta \gamma_\beta^0 \end{Bmatrix} = 0 \tag{26}$$

Where [SR]:

$$\begin{bmatrix} [SR^\alpha] & 0 & 0 & 0 & 0 \\ 0 & [SR^\beta] & 0 & 0 & 0 \\ 0 & 0 & [SR^\zeta] & 0 & 0 \\ 0 & 0 & 0 & [SR^{\gamma^\alpha}] & 0 \\ 0 & 0 & 0 & 0 & [SR^{\gamma^\beta}] \end{bmatrix} \tag{27}$$

Notice that $[SR^\alpha], \dots, [SR^{\gamma^\beta}]$ have different dimensions.

The terms MR_{ij} and SR_{ij} are spatial derivatives of the stiffness properties obtained after integrating by parts the derivatives of the displacement d.o.f. So Eqs. (24) and (26) represent partial differential equations. The equations obtained for the strain energy contribution in the plane (β, ζ) and for the membrane contributions are similar.

As mentioned above, just remind that the equations involving in-plane energy are disregarded, since laminated and sandwich composites have weak out-of-plane properties and strong in-plane properties. Thus, just the lines that multiply $\delta u_\zeta^0, \delta \gamma_\alpha^0, \delta \gamma_\beta^0$ require a simultaneous solution. As the extremal condition under variation of the stiffness properties yields irrespective of the response, equations (24) and (26) split into a set of partial differential equations of the following form:

$$\sum MR_{3j} = 0; \sum MR_{4j} = 0; \sum MR_{5j} = 0 \tag{28}$$

$$\sum SR_{3j} = 0; \sum SR_{4j} = 0; \sum SR_{5j} = 0 \tag{29}$$

The distribution of the spatial stiffness properties that make extremal the bending and transverse shear energy contributions is obtained by the simultaneous solution of Eqs. (28) and (29). The law of variation of the fibre orientation angle is obtained from the distribution of $Q_{ij}(\alpha, \beta)$ (ij= 11, 12, 13, 16, 22, 23, 26, 44, 45, 55, 66); this distribution is obtained in a straightforward way since the Q_{ij} integrated across the thickness give the stiffness terms.

The tailoring optimisation of the face plies looks for the optimal stiffness distributions in (α, β) ; for this reason the spatial derivatives of the stiffness coefficients in α, β determine the form of solutions, while the derivatives in ζ are irrelevant. On the contrary, as the core properties are assumed to vary across the thickness, the form of solutions is determined by the spatial derivatives in ζ .

In this section, in order to arrive at the solution for MR_{ij} and SR_{ij} , the terms that appear in the strain component $\varepsilon_{\alpha\alpha}$ and $\varepsilon_{\alpha\zeta}$ are discussed. Nevertheless the conclusions hold also for $\varepsilon_{\alpha\beta}$, $\varepsilon_{\beta\beta}$ and $\varepsilon_{\beta\zeta}$, since they have the same mathematical features. It could be noticed that in C_α^i , C_β^i , D_β^i , D_α^i , as well as in the other high order coefficients and in b^i to e^i and finally in the continuity functions Φ_α^k , Φ_β^k , Ψ^k , Ω^k , C_u^k , C_v^k , C_w^k , the in-plane derivatives of various orders of the functional d.o.f. play an important role, so determine the form of solutions. In these terms powers of ζ of various orders are involved, which generate stiffness coefficients that are irrelevant for the in-plane optimisation process. For example, this is the case of terms ζ/R_α and ζ/R_β that appear in the strain definitions and in the expressions of H_α and H_β .

On the contrary, the order of differentiation of the stiffness quantities those appear in Eqs. (28) and (29) is determined by the derivatives of the displacements explicitly appearing in the strains and those implicitly appearing in C_α^i ,

C_β^i , D_β^i , D_α^i , as well as in the other high order coefficients and in b^i to e^i and in the continuity functions Φ_α^k , Φ_β^k , Ψ^k , Ω^k , C_u^k , C_v^k , C_w^k . In fact, once the displacement fields are substituted and the strain energy is computed, the derivatives of the d.o.f. are integrated by parts, resulting into derivatives of the stiffness quantities.

To show the contributions that are obtained according to the scheme described previously, here the strain energy due to the transverse shear strain $\varepsilon_{\alpha\zeta}$ is written in explicit form.

Similar considerations are also valid for $\varepsilon_{\alpha\alpha}$ and $\varepsilon_{\beta\zeta}$ as mentioned above. The following expression derives from the strain energy due to $\varepsilon_{\alpha\zeta}$:

$$\int_{\Omega} \int_{\zeta} Q_{44} \left[\frac{H_\alpha}{H_\zeta} \left(\frac{u_\alpha}{H_\alpha} \right)_{,\zeta} + \frac{H_\zeta}{H_\alpha} \left(\frac{u_\zeta}{H_\zeta} \right)_{,\alpha} \right]^2 d\zeta d\Omega + \int_{\Omega} \int_{\zeta} Q_{45} \left\{ \left[\frac{H_\alpha}{H_\zeta} \left(\frac{u_\alpha}{H_\alpha} \right)_{,\zeta} + \frac{H_\zeta}{H_\alpha} \left(\frac{u_\zeta}{H_\zeta} \right)_{,\alpha} \right] \left[\frac{H_\beta}{H_\zeta} \left(\frac{u_\beta}{H_\beta} \right)_{,\zeta} + \frac{H_\zeta}{H_\beta} \left(\frac{u_\zeta}{H_\zeta} \right)_{,\beta} \right] \right\} d\zeta d\Omega \quad (30)$$

The expressions of the displacement in the strain components were left in implicit form for limiting the length of the formula. The derivatives of displacements are converted integrating by parts obtaining the following expressions of the stiffness coefficients

$$\int_{\zeta} \left\{ Q_{44} \left(\frac{H_\alpha}{H_\zeta} \right) \left(1 + \frac{\zeta}{R_\alpha} \right) \right\}_{,\zeta}^2 d\zeta + \int_{\zeta} \left\{ Q_{44} \left(\frac{H_\alpha}{H_\zeta} \right) \frac{\zeta}{A_\alpha} \right\}_{,\zeta}^2 d\zeta + \int_{\zeta} \left\{ Q_{44} \left(\frac{H_\alpha}{H_\zeta} \right) \left(1 + \frac{\zeta}{R_\alpha} \right) \frac{\zeta}{A_\alpha} \right\}_{,\zeta} d\zeta \quad (31)$$

Similar expressions are obtained integrating by parts the derivatives of the d.o.f. that appear in the expressions of the displacements written in implicit form in (30). The stiffness quantities related to the contribution by $\varepsilon_{\alpha\alpha}$ to the strain energy are obtained in a similar way:

$$\int_{\Omega} \int_{\zeta} Q_{11} \left[\frac{u_{\alpha,\alpha}}{H_\alpha} + \frac{H_{\alpha,\beta}}{H_\alpha H_\beta} u_\beta + \frac{H_{\alpha,\zeta}}{H_\alpha H_\zeta} u_\zeta \right]^2 d\zeta d\Omega + \int_{\Omega} \int_{\zeta} Q_{12} \left\{ \left[\frac{u_{\alpha,\alpha}}{H_\alpha} + \frac{H_{\alpha,\beta}}{H_\alpha H_\beta} u_\beta + \frac{H_{\alpha,\zeta}}{H_\alpha H_\zeta} u_\zeta \right] \left[\frac{u_{\beta,\beta}}{H_\beta} + \frac{H_{\beta,\alpha}}{H_\alpha H_\beta} u_\alpha + \frac{H_{\beta,\zeta}}{H_\beta H_\zeta} u_\zeta \right] \right\} d\zeta d\Omega$$

$$\begin{aligned}
& + \int_{\Omega_{\zeta}} Q_{13} \left\{ \frac{u_{\alpha,\alpha}}{H_{\alpha}} + \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} u_{\beta} + \frac{H_{\alpha,\zeta}}{H_{\alpha}H_{\zeta}} u_{\zeta} \right\} \cdot \frac{u_{\zeta,\zeta}}{H_{\zeta}} \Bigg\} d\zeta d\Omega + \\
& + \int_{\Omega_{\zeta}} Q_{16} \left\{ \frac{u_{\alpha,\alpha}}{H_{\alpha}} + \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} u_{\beta} + \frac{H_{\alpha,\zeta}}{H_{\alpha}H_{\zeta}} u_{\zeta} \right\} \cdot \\
& \left[\frac{H_{\alpha}}{H_{\beta}} \left(\frac{u_{\alpha}}{H_{\alpha}} \right)_{,\beta} + \frac{H_{\beta}}{H_{\alpha}} \left(\frac{u_{\beta}}{H_{\beta}} \right)_{,\alpha} \right] d\zeta d\Omega
\end{aligned} \quad (32)$$

So the stiffness coefficients result:

$$\begin{aligned}
& \int_{\zeta} \left\{ \frac{1}{H_{\alpha}} \left(1 + \frac{\zeta}{R_{\alpha}} \right) \right\}_{,\alpha}^2 d\zeta \\
& \int_{\zeta} \left\{ \frac{1}{H_{\alpha}} \frac{\zeta}{A_{\alpha}} \right\}_{,\alpha}^2 d\zeta \\
& \int_{\zeta} \left\{ \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \left(1 + \frac{\zeta}{R_{\beta}} \right) \right\}^2 d\zeta \\
& \int_{\zeta} \left\{ \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \frac{\zeta}{A_{\beta}} \right\}^2 d\zeta \\
& \int_{\zeta} \left\{ \left(\frac{1}{H_{\alpha}} \right)^2 \left(1 + \frac{\zeta}{R_{\alpha}} \right) \frac{\zeta}{A_{\alpha}} \right\}_{,\alpha} d\zeta \\
& \int_{\zeta} \left\{ \frac{1}{H_{\alpha}} \left(1 + \frac{\zeta}{R_{\alpha}} \right) \right\}_{,\alpha} \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \left(1 + \frac{\zeta}{R_{\beta}} \right) d\zeta \\
& \int_{\zeta} \left\{ \frac{1}{H_{\alpha}} \frac{\zeta}{A_{\alpha}} \right\}_{,\alpha} \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \left(1 + \frac{\zeta}{R_{\beta}} \right) d\zeta \\
& \int_{\zeta} \left\{ \frac{1}{H_{\alpha}} \frac{\zeta}{A_{\alpha}} \right\}_{,\alpha} \frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \frac{\zeta}{A_{\beta}} d\zeta \\
& \int_{\zeta} \left[\frac{H_{\alpha,\beta}}{H_{\alpha}H_{\beta}} \right]^2 \left(1 + \frac{\zeta}{R_{\beta}} \right) \frac{\zeta}{A_{\beta}} d\zeta
\end{aligned} \quad (33)$$

As it can be noticed from Eq. (33), in this case are involved either non derived terms, derivatives of the functional d.o.f. across the thickness as well as in-plane derivatives, which play when the stiffness properties of the core or those of faces are optimised, respectively.

Because the radii R_{α} , R_{β} are variable from point to point, the solution also depends by their variation. However when treating cylindrical, conical, or spherical shells it is possible to assume that the radii remain constants like; this enables to obtain a closed form solution.

The extremal conditions for the strain energy under variation of the stiffness properties determine a set of partial differential equations up to the fourth order in (α, β) and up to the third order in ζ , since the continuity functions involve derivatives of displacements up to the third order. This set of equations has to be solved simultaneously, and the solutions of the optimization problem of the face plies are represented by exponential functions for elastic coefficients with indices 11, 12, 13, 16, 22, 23, 26, 36, 44, 45, 55, 66 and constant for elastic coefficients with indices 44, 45, 55, since just first order derivatives of the stiffness properties with these indices are involved. This is physically consistent since the orientation of the fibres over the plane (α, β) does not affect the transverse shear properties of the plies.

Therefore the solution of the extremal problem of the face plies is

$$\begin{aligned}
Q_{ij} &= \sum_{p=1}^P \left[A_{1p}^{ij} e^{(p\alpha + \phi_n^{\alpha})} + k_{\alpha} \right] \cdot \left[A_{2p}^{ij} e^{(p\beta + \phi_n^{\beta})} + k_{\beta} \right] \\
Q_{44} &= L \\
Q_{45} &= M \\
Q_{55} &= N
\end{aligned} \quad (34)$$

The appropriate amplitude, phase, mean value and period in (34) are determined enforcing conditions such as the stiffness at the bounds of the domain and a convex or a concave shape, in order to determine whether the solution minimises or maximises the strain energy components.

The solution has to be physically consistent, so it is necessary to enforce thermodynamic constraint conditions, conservation of energy, Lempriere's and Chentsov's conditions.

The strain energy due to transverse shears contains non derived terms and derivatives up to the third order in ζ , therefore also the core properties across the thickness should be distributed as exponential functions:

$$Q_{ij} = \sum_{p=1}^P A_{1p}^{ij} e^{(p\zeta + \phi_n^{\zeta})} + A_{2p}^{ij} e^{(-p\zeta + \phi_n^{\zeta})} + k_{\zeta} \quad (35)$$

In the next section, the applications of this optimization technique will be numerically assessed for laminated and sandwich plates as well as for sandwich spherical panels.

4 NUMERICAL APPLICATIONS

First of all two optimized orientation of the fibre are determined according to the solutions of Eq. (34). Therein the solution is represented by exponential functions, in the current cases these forms are approximated with polynomial expansion, whose order is variable:

$$\begin{aligned}
 Q_{ij} &= \sum_{g=1}^G A_g \alpha^g + B_g \beta^g \\
 Q_{44} &= L \\
 Q_{45} &= M \\
 Q_{55} &= N
 \end{aligned} \tag{36}$$

The order of expansion G is determined considering accuracy as well as computational effort.

The polynomial expansion makes easier the algebraic manipulation and is suited for the actual manufacturing technologies.

As for Eq. (34) the unknown coefficient are determined by imposing thermodynamic condition and condition such as the stiffness at the bounds of the domain, a convex or a concave shape and other condition till saturation of all the index. It is important to underline that the mean properties (i.e. the mean value of the functions of their powers of various order) of the optimized ply are the same of that unoptimized.

Two distributions of the fibres are considered: one that minimizes the shear component of the strain energy (here indicate with Min Shear), one that minimizes the bending component of the strain energy (Min Bending).

From the practical point of view a Min Bending distribution is obtained imposing maximum bending stiffness at the centre of the shell and minimum at the edge.

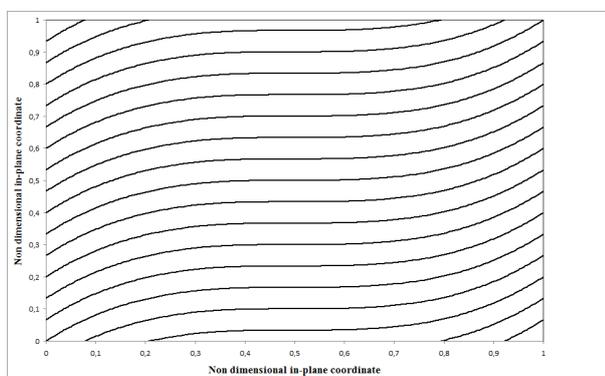


Figure 1 Fibre distribution for a Min Bending ply.

A Min Shear distribution is obtained imposing maximum bending stiffness at the edge of the shell and minimum at the centre.

The two optimized orientation of the fibre here calculated have been applied to laminated plate and to a sandwich spherical panel.

A simply-supported laminated square plate under sinusoidal loading, a simply-supported sandwich square plate under sinusoidal loading and a simply supported spherical sandwich panel under sinusoidal loading are considered in the numerical applications.

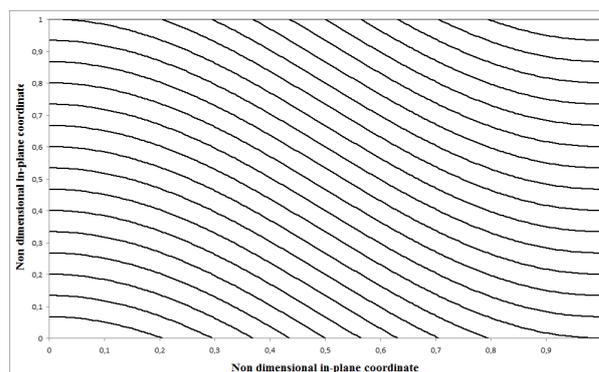


Figure 2 Fibre distribution for a Min Shear ply.

The solution is calculated as a trigonometric series expansion that enables to fulfil the boundary conditions:

$$u_{\alpha}^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N A_{mn} \cos\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \sin\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{37}$$

$$u_{\beta}^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N B_{mn} \sin\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \cos\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{38}$$

$$u_{\zeta}^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N C_{mn} \sin\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \sin\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{39}$$

$$\gamma_{\alpha}^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N D_{mn} \cos\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \sin\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{40}$$

$$\gamma_{\beta}^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N E_{mn} \sin\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \cos\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{41}$$

The distributed loading is represented as:

$$p^0(\alpha, \beta) = \sum_{m=1}^M \sum_{n=1}^N P_{mn} \sin\left(\frac{m\pi}{\psi_{\alpha}} \alpha\right) \sin\left(\frac{n\pi}{\psi_{\beta}} \beta\right) \tag{42}$$

In all the examined cases $M=N=1$ is sufficient for obtaining accurate results.

Substituting the previous approximate representation into the governing equations, a set of linear algebraic equations is obtained by the Rayleigh-Ritz method that enables to determine the unknown amplitudes.

4.1 LAMINATED AND SANDWICH PLATES

A simply-supported laminated square plate under sinusoidal loading is considered.

Even though not realistic, this case of analysis is the only one that presents an exact three-dimensional elasticity solution, as shown by Pagano [54].

According with Pagano the plate is characterized by a thickness ratio $S=L/h$ ($L=L_x$) of 4, though with no practical application it is a severe test for the model. It could be noticed that analyzing the gain at various S , this parameter linearly rises till $S=100$, then it assumes a constant value.

The displacements and stresses are reported where they assume their maximum, in the following normalised form:

$$\begin{aligned} \bar{\sigma}_{xz} &= \frac{\sigma_{xz}\left(0, \frac{L_y}{2}, z\right)}{p^0 S} & \bar{\sigma}_{yz} &= \frac{\sigma_{yz}\left(\frac{L_x}{2}, 0, z\right)}{p^0 S} \\ \bar{u} &= \frac{E_T u\left(0, \frac{L_y}{2}, z\right)}{hp^0 S^3} & \bar{w} &= 100 \frac{E_T w\left(\frac{L_x}{2}, \frac{L_y}{2}, z\right)}{hp^0 S^4} \end{aligned} \quad (43)$$

the un-optimized material has the following properties: $E_L/E_T=25$; $G_{LT}/E_T=0.5$; $G_{TT}/E_T=0.2$; $\nu_{LT}=0.25$; while the lay-ups considered are reported in Figure 3.

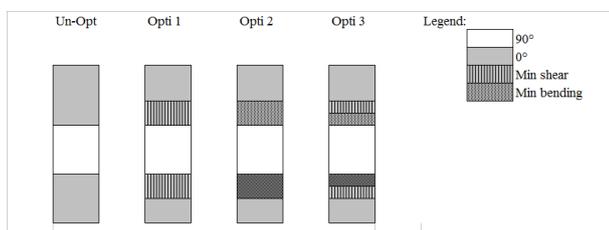


Figure 3 Lay-ups considered.

Figure 5 represents the through-the-thickness variation of the shear stress $\bar{\sigma}_{xz}$ of the lay-ups considered while Figure 6 represents the through-the-thickness variation of the shear stress $\bar{\sigma}_{yz}$.

In Table I is reported the deflection in the centre of the plates for the considered lay-ups.

Table I - Deflection in the centre of the plates

	Un Opt	Opti 1	Opti 2	Opti 3
\bar{w}	1,954	1,919	1,903	1,890
gain		-1,80%	-2,60%	-3,30%

From the reported results it can be noticed the effects of the optimization. Remembering that the aim of the optimization is the reduction of the interlaminar stresses, from this viewpoint the most efficient lay-up Opti 1, whose interlaminar shear stress sensibly decreases respect to the un-optimized case.

Even if, in order to contain the length of the paper, the in-plane stress are not reported, it could be underlined that in the optimized configurations these stresses rise up respect to the un-optimized one. However this is not a problem since the strength of the material still remains higher than the new stresses, so the laminate is far from failure.

To better underline the advantage of the optimization technique, a simply-supported sandwich square plate under sinusoidal loading is considered.

The plate is still characterized by a thickness ratio S of 4, the displacements and stresses are reported where they assume their maximum, in the following normalised form:

$$\bar{\sigma}_{xz} = \frac{\sigma_{xz}\left(0, \frac{L_y}{2}, z\right)}{p^0} \quad \bar{w} = \frac{w\left(\frac{L_x}{2}, \frac{L_y}{2}, z\right)}{p^0 h} \quad (44)$$

The face layers are made of materials 1 to 3, while the core is made of material 4, whose un-optimized mechanical properties are as follows. MAT 1: $E_1=E_3=1$ GPa, $G_{13}=0.2$ GPa, $\nu_{13}=0.25$; MAT 2: $E_1=33$ GPa, $E_3=1$ GPa, $G_{13}=0.8$ GPa, $\nu_{13}=0.25$; MAT 3: $E_1=25$ GPa, $E_3=1$ GPa, $G_{13}=0.5$ GPa, $\nu_{13}=0.25$; MAT 4: $E_1=E_3=0.05$ GPa, $G_{13}=0.0217$ GPa, $\nu_{13}=0.15$. MAT 1 is rather weaker in tension-compression and shear compared to MAT 2, whereas MAT 3 is stiff in tension-compression but rather weak in shear. As usual for core materials, MAT 4 compared to the other materials is very weak in tension-compression and rather weak in shear. According to what previously discussed, the sandwich plate is simulated as a multilayered plate and the lay-ups considered are:

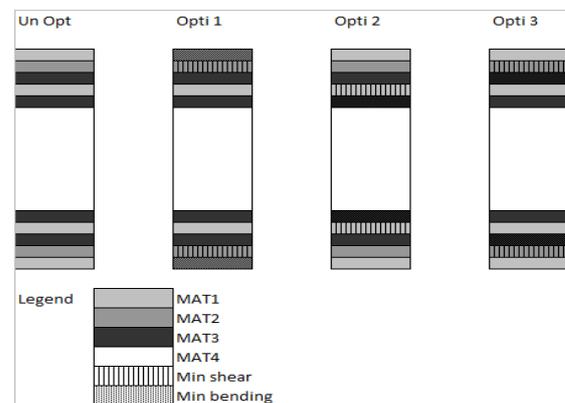


Figure 4 Lay-ups considered.

Figure 7 represents the through-the-thickness variation of the shear stress $\bar{\sigma}_{xz}$ of the lay-ups considered while in Table II and Table III the deflection and the interlaminar shear stress are reported.

Table II - Deflection for the lay-ups considered

	Un Opt	Opti 1	Opti 2	Opti 3
\bar{w}_{max}	$1,78 \cdot 10^{-8}$	$1,44 \cdot 10^{-8}$	$1,52 \cdot 10^{-8}$	$1,44 \cdot 10^{-8}$
gain		-24,10%	-17,14%	-24,08%

Table III - Interlaminar shear stress for the lay-ups considered

	Un Opt	Opti 1	Opti 2	Opti 3
$\bar{\sigma}_{xz,face}$	0,78	0,41	0,50	0,49
gain		-91,28%	-58,13%	-59,23%
$\bar{\sigma}_{xz,core}$	0,12	0,080	0,094	0,080
gain		-53,73%	-31,08%	-54,35%

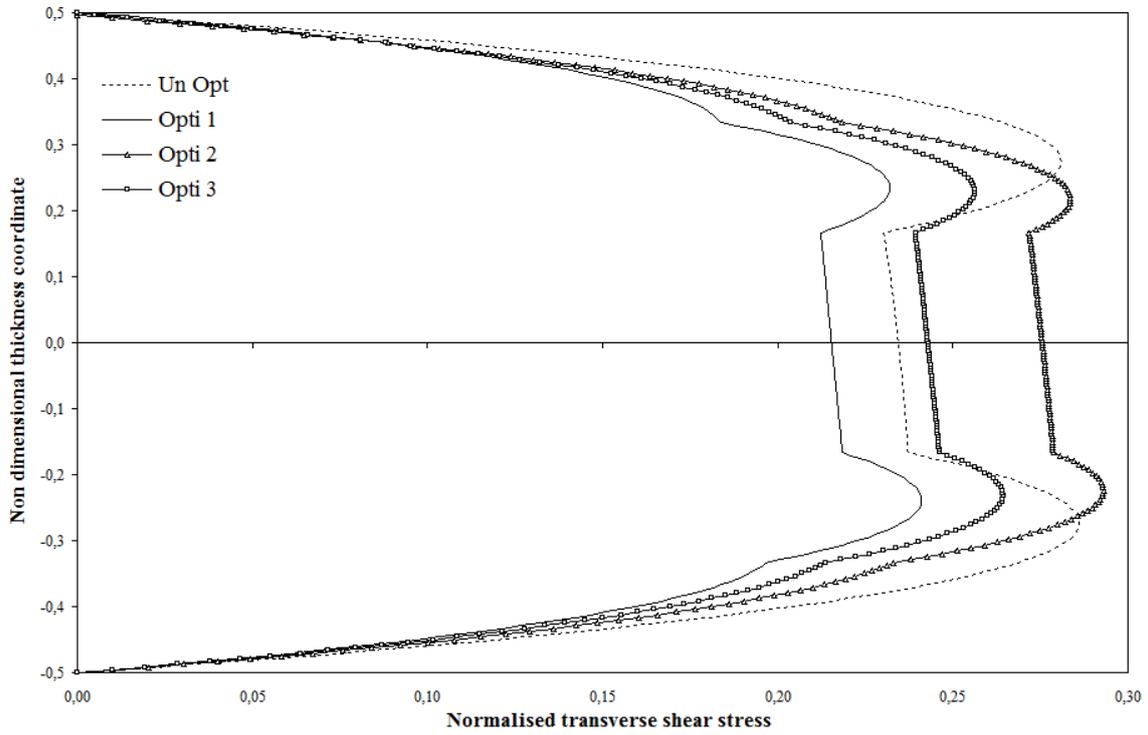


Figure 5 Through-the-thickness distribution of the transverse shear stress $\bar{\sigma}_{xz}$ (laminated plate).

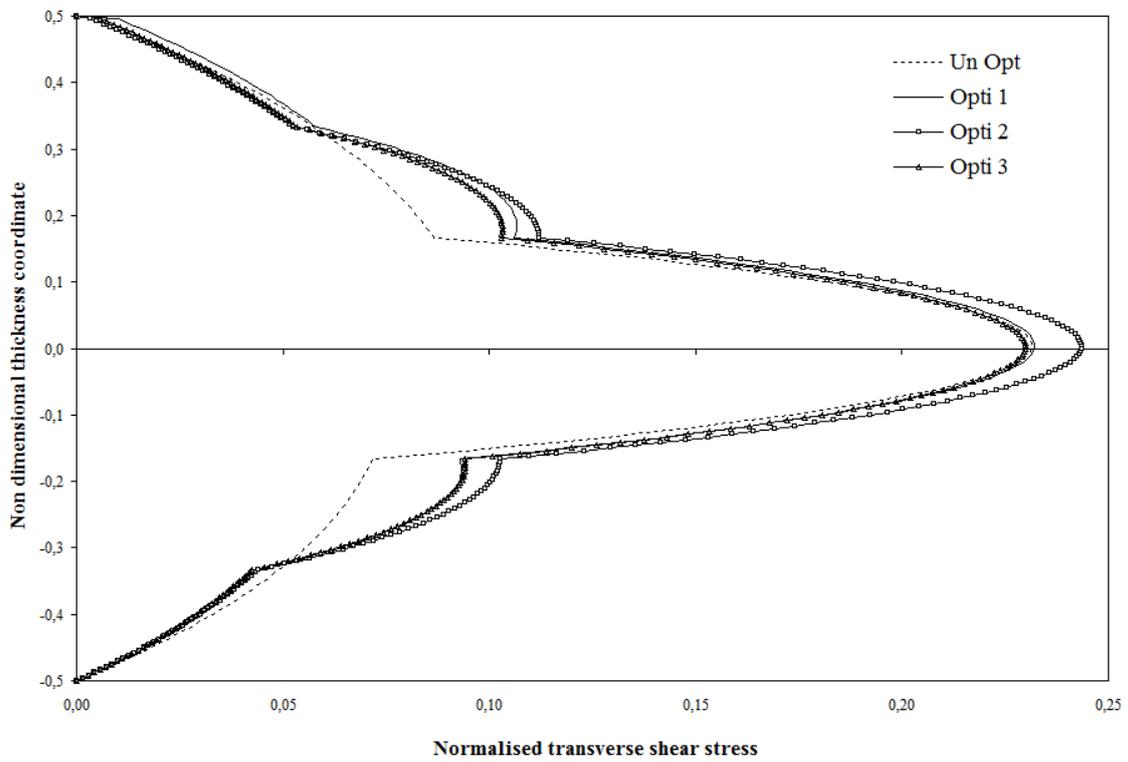


Figure 6 Through-the-thickness distribution of the transverse shear stress $\bar{\sigma}_{yz}$ (laminated plate).

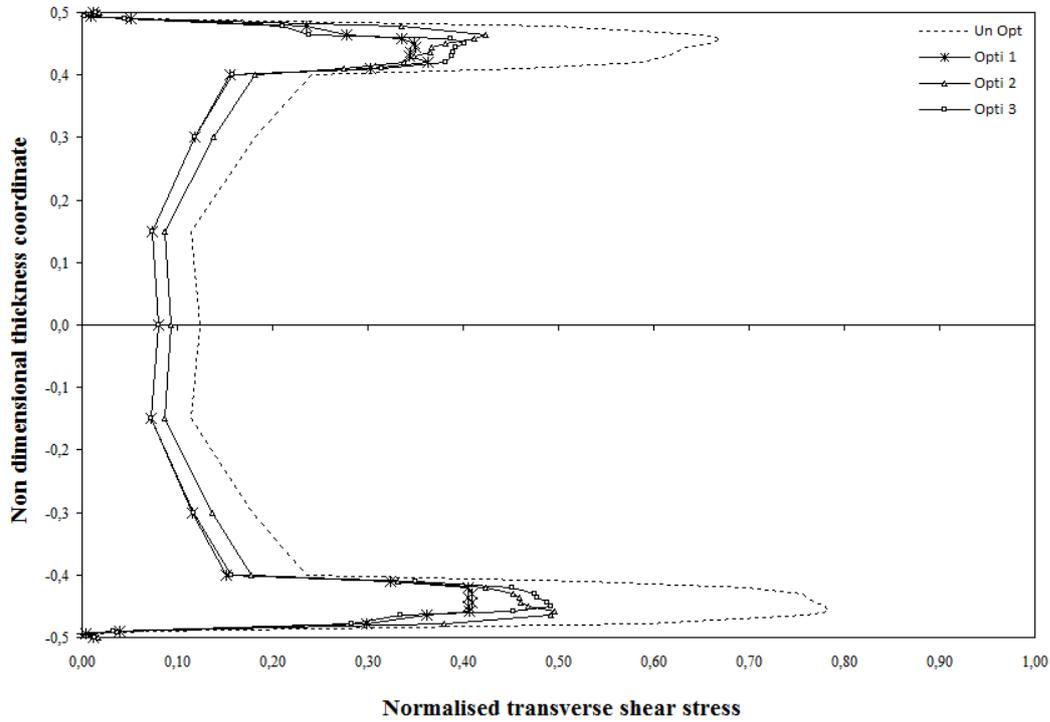


Figure 7 Through-the-thickness distribution of the transverse shear stress $\bar{\sigma}_{xz}$ (sandwich plate).

4.2 SPHERICAL PANEL

Here a simply-supported spherical sandwich panel under sinusoidal loading is considered.

The panel has $L_\alpha = L_\beta$, $R_\alpha = R_\beta = 100$ and is characterized by a radius to thickness ratio of 4. The angle subtended by the panel are $\psi_\alpha = \psi_\beta = \pi/3$.

The displacements and stresses are reported where they assume their maximum, in the following normalised form:

$$\bar{\sigma}_{\alpha\zeta} = \frac{\sigma_{\alpha\zeta}\left(0, \frac{\psi_\beta}{2}, \zeta\right)}{p^0}, \quad \bar{u}_\alpha = \frac{u_\alpha\left(0, \frac{\psi_\beta}{2}, \zeta\right)}{p^0 h}, \quad \bar{u}_\zeta = \frac{u_\zeta\left(\frac{\psi_\alpha}{2}, \frac{\psi_\beta}{2}, \zeta\right)}{p^0 h} \quad (45)$$

The materials are the same already used for the sandwich plates and the lay-ups considered are represented in Figure 8.

Figure 9 represents the through-the-thickness variation of the shear stress $\bar{\sigma}_{\alpha\zeta}$ of the lay-ups considered while Figures 10 and 11 represent respectively the through-the-thickness variation of the displacements \bar{u}_α and \bar{u}_ζ .

The deflection and the interlaminar shear stress are reported in Table IV and in Table V, respectively.

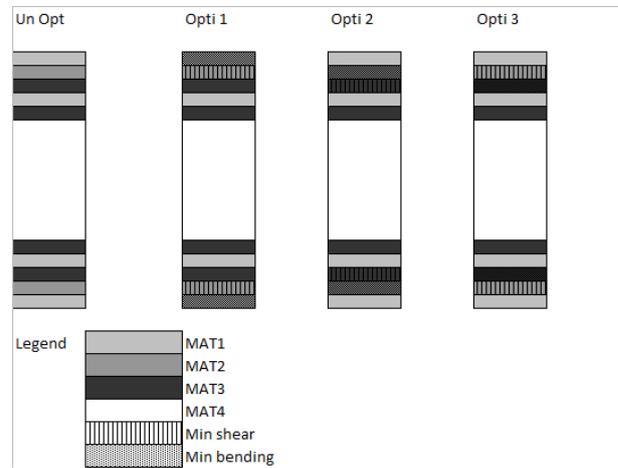


Figure 8 Lay-ups considered.

Table IV - Deflection for the lay-ups considered

	Un Opt	Opti 1	Opti 2	Opti 3
$\bar{u}_{\zeta \max}$	$1,83 \cdot 10^{-8}$	$1,50 \cdot 10^{-8}$	$1,57 \cdot 10^{-8}$	$1,50 \cdot 10^{-8}$
gain		-21,87%	-16,64%	-21,97%

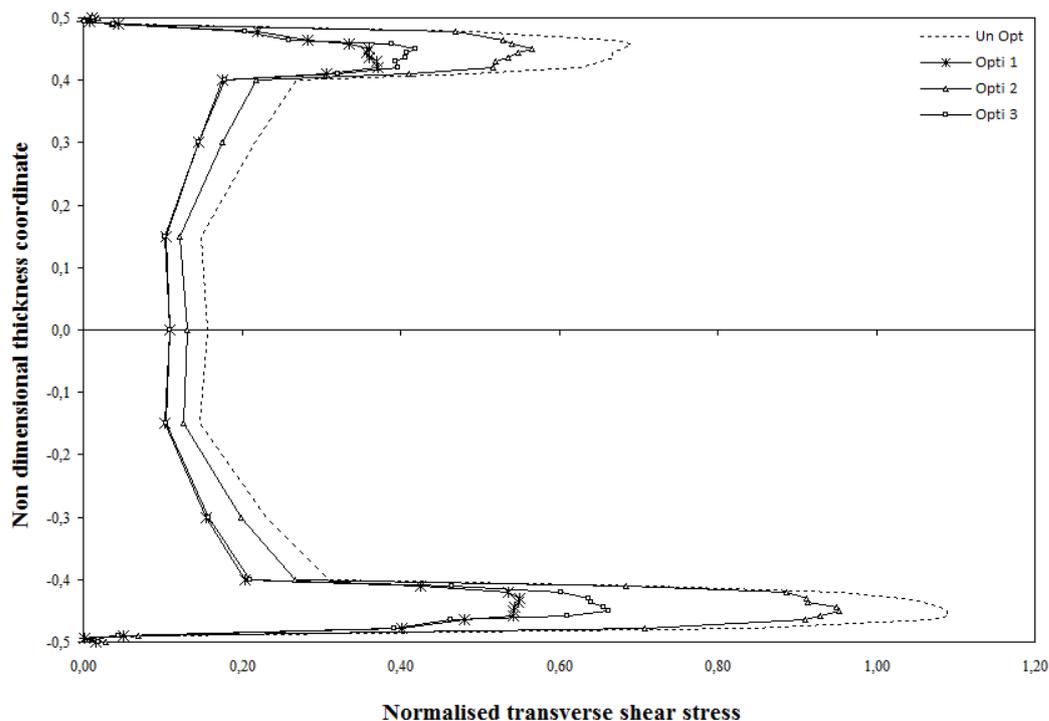
Table V - Interlaminar shear stress for the lay-ups considered

	Un Opt	Opti 1	Opti 2	Opti 3
$\overline{\sigma}_{xz, face}$	1,09	0,55	0,95	0,66
gain		-	-14,45%	-64,94%
$\overline{\sigma}_{xz, core}$	0,16	0,1098	0,1309	0,1091
gain		-42,22%	-19,31%	-43,12%

The numerical results show that the presence of the curvature enhances the coupling effect between in-plane stress and shear stress. This determines an unsymmetrical distribution of the shear stress $\overline{\sigma}_{\alpha\zeta}$. The advantages of the optimization are still evident. In particular the deflection as well as the interlaminar stress sensibly decrease. Many stress based criteria have been applied in order to verify whether the coupling effects determine a failure on the panel. It was found that all the criteria show results in a well agreement. In fact, even if the optimization process determines the rise of in-plane stresses, a consistent reduction of the delamination index is achieved, with beneficial effect on strength at the onset of delamination. In particular, considering the Hou-Petricin-Ruiz's [55] criterion, it can be obtained a consistent decrease of the failure index for delamination up to 80%, considering the combined action of in-plane and out-of-plane stresses in the most critical region close to the interfaces.

5 CONCLUDING REMARKS

A tailoring optimization technique recently developed has been extended to sandwich shells, with the aim to minimize the strain energy due to out-of-plane shear strains and to maximize that due to their bending counterparts. The solution of the Euler-Lagrange equations of an extremal problem, where the stiffness properties are assumed as the master field, enables to determine the in-plane distribution of the face ply properties and the through-the-thickness distribution of the core properties. The optimal distribution of stiffness properties has been approximated with a polynomial variation that can be easily made with the current manufacturing technologies. Applications were presented to laminated and sandwich plates and also to spherical sandwich shell. As shown by the numerical results the critical interlaminar stress concentration at the interfaces is consistently reduced from 10 to 40 % without any bending stiffness loss, since the transverse displacement in all the examined cases decreases. The optimization technique is more effective for sandwich plate and shell, than for their laminated counterparts. It could be noticed that the curvature enhances the coupling effect between in-plane stress and shear stress, determining an unsymmetrical distribution of the shear stress $\overline{\sigma}_{\alpha\zeta}$. Thanks to the insertion of optimized ply in laminated and sandwich panels, the failure index for tensile/compressive failure of fibres and matrix decreases up to 80%. Therefore it could be underlined that the tailoring optimization transfers energy from out-of-plane to in-plane modes.

Figure 9 Through-the-thickness distribution of the transverse shear stress $\overline{\sigma}_{\alpha\zeta}$.

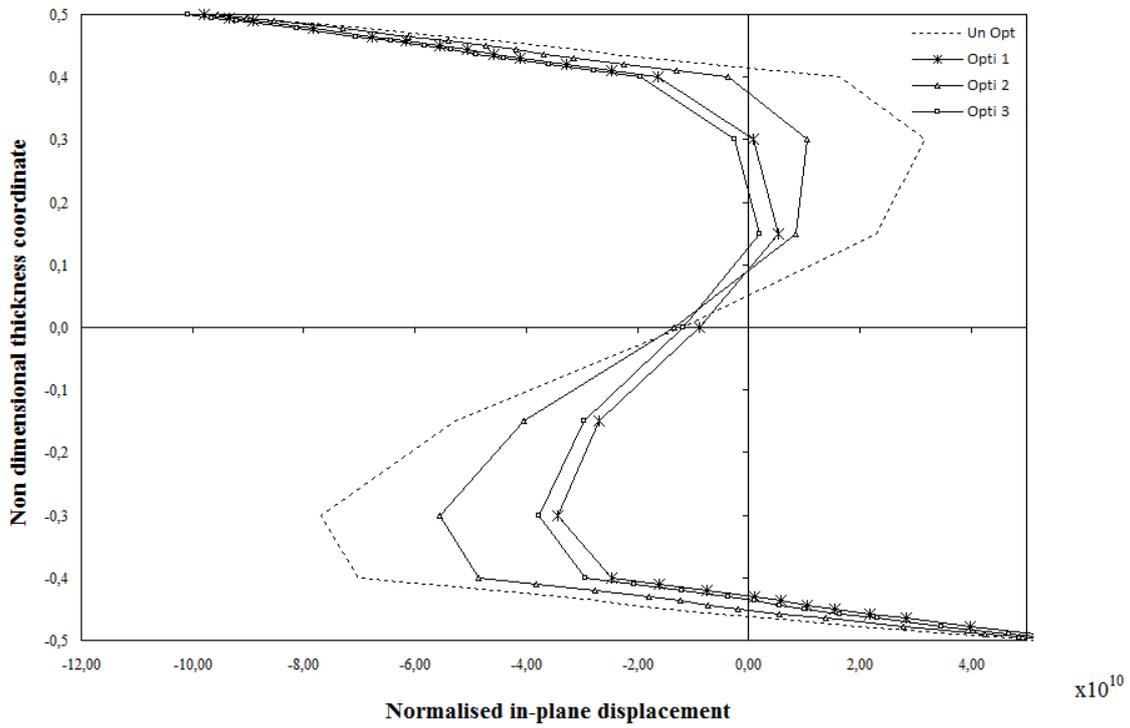


Figure 10 Through-the-thickness distribution of the in-plane displacement \bar{u}_α .

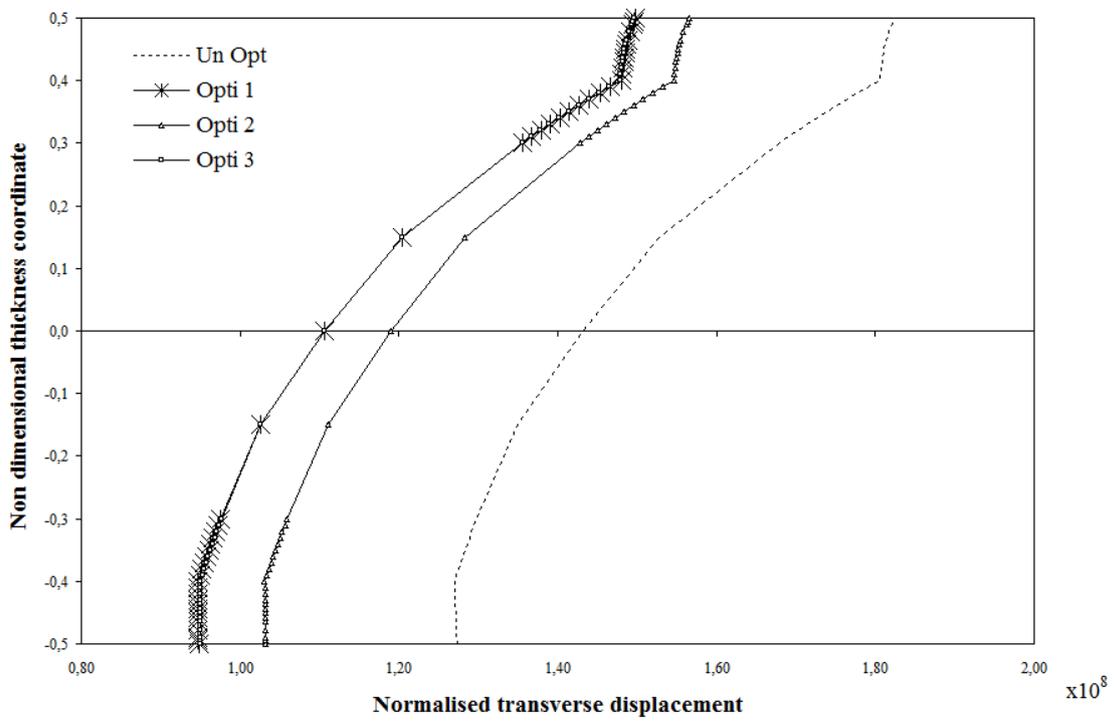


Figure 11 Through-the-thickness distribution of the in-plane displacement \bar{u}_ζ .

REFERENCES

- [1] Rowland R.E., Strength (failure) theories and their experimental correlations. Handbook of Composites, Vol. 3, pp. 71–125 in, Elsevier Science Publ., 1985.
- [2] Tennyson R.C. and Wharam G.E., Evaluation of failure criterion for graphite-epoxy, Technical report NASA CR-172547, NASA, 1985.
- [3] Nahas M. N., Survey of failure and post-failure theories of laminated fibre reinforced composites. J. Compos. Tech. Res., Vol. 8, pp. 138–153, 1986.
- [4] Bolotin V.V., Delaminations in composite structures: its origin, buckling, growth and stability. Compos. B: Eng, Vol. 27, pp. 129–145, 1996.
- [5] Echaabi J.F. and Trochu F., Review of failure criteria of fibrous composite materials. Polym. Compos., Vol. 17, pp. 786–798, 1996.
- [6] Paris F., A study of failure criteria of fibrous composite materials, Technical report NASA/CR-2001-210661, NASA, 2001.
- [7] Icardi U., Locatto S. and Longo A., Assessment of recent theories for predicting failure of composite laminates. Appl. Mech. Rev., Vol. 60, No. 2, pp. 76–86, 2007.
- [8] Chaperon P., Jones R., Heller S., Pitt F., Rose A., A methodology for structural optimization with damage tolerance constraints. Eng Failure Analysis, Vol. 7, pp. 281–300, 2000.
- [9] Jones R., Peng D., Chaperon P., Pitt S., Abramson D., Peachey T., Structural optimization with damage tolerant constraints. Theoretical & Applied Fracture Mechanics, Vol. 43, pp. 133–155, 2000.
- [10] Keller D., Optimization of ply angles in laminated composite structures by a hybrid asynchronous, parallel evolutionary algorithm. Composite Structures, Vol. 92, pp. 2781–2790, 2010.
- [11] Kaye R., Heller M., Investigation of Shape optimization for the design of life extension options for an F/A-18 airframe FS 470 bulkhead. J. Strain Analysis, Vol. 35, pp. 493–505, 2000.
- [12] Sivakumar K., Iyengar N.G.R., Kalyanmoy D., Optimization of composite laminates with cutouts using genetic algorithm, variable metric and complex search methods, Eng Opt, Vol. 32, pp. 635–657, 2000.
- [13] Uys P.E. and Sankar B.V., Optimal design of a hoist structure frame. Applied Math Modelling, Vol. 27, pp. 963–982, 2003.
- [14] Krishnapillai K. and Jones R., Three-dimensional structural design optimization based on fatigue implementing a genetic algorithm and a non-similitude crack growth law. Finite Elements in Analysis and Design, Vol. 45, pp. 132–146, 2009.
- [15] Hyer M.W. and Charette R.F., Use of curvilinear fiber format in composite structure design. AIAA J, Vol. 29, No. 6, pp. 1011–1115, 1991.
- [16] Jegley D., Tatting B. and Gürdal Z., Tow-steered panels with holes subjected to compression or shear loading. Proc. of AIAA/ASME/ASCE/AHS/ASC 46th Structures, Structural Dynamics and Materials Conference, Austin TX (USA), AIAA 2005-2017,2005.
- [17] Lopes C.S., Gürdal Z. and Camanho P.P., Variable-stiffness panels: buckling and first-ply failure improvements over straight-fibre laminates. Computers & Struct, Vol. 86, pp. 897–907, 2008.
- [18] Pedersen P., A note on design of fiber-nets for maximum stiffness. J. Elasticity, Vol. 73, No. 1–3, pp. 127–145, 2003.
- [19] Setoodeh S., Abdalla M.M., Gürdal Z. and Tatting B., Design of variable-stiffness composite laminates for maximum in-plane stiffness using lamination parameters. Proceeding of 46th AIAA/ASME/ASCE/AHS/ASC Struc., Struct. Dynam. and Appl. Conf., 13th AIAA/ ASME/ AHS Adap. Struc. Conf., 7th AIAA Non-Determ. Appr. Forum, Austin TX (USA), pp. 3473–3481, 2005.
- [20] Diaconu C.G., Sato M. and Semine H., Lay-up optimization of symmetrically laminated thick plates for fundamental frequencies using lamination parameters. Struct Multidisciplinary Optim., Vol. 24, pp. 302–311, 2002.
- [21] Narita Y. and Hodgkinson J.M. Layerwise optimisation for maximising the fundamental frequencies of point-supported rectangular laminated composite plates. Compos Struct, Vol. 69, No. 2, pp. 127–135, 2005.
- [22] Abdalla M.M., Setoodeh S. and Gürdal Z., Design of variable stiffness composite panels for maximum fundamental frequency using lamination parameters. Composite Structures, Vol. 8, pp. 283–291, 2007.
- [23] Ghiasi H., Fayazbakhsh K., Pasini D., Lessard L., Optimum stacking sequence design of composite materials Part II: variable stiffness design. Composite Structures, Vol. 93, pp. 1–13, 2010.
- [24] Javidrad D., Nouri R., A simulated annealing method for design of laminates with required stiffness properties. Composite Structures, Vol. 93, pp. 1127–1135, 2011.
- [25] Jung W.Y.A., Combined honeycomb and Solid Viscoelastic Material for Structural Damping Applications. Proc. of Thrust Area 2: Seismic Retrofit of Acute Care Facilities. Buffalo, pp. 41–43, 2001.
- [26] Lakes R.S., High damping composite materials: effect of structural hierarchy. J. Composite Materials, Vol. 36, No. 3, pp. 287–297, 2002.
- [27] Barth J., Fabrication of complex composite structures using advanced fiber placement technology. Proc. of 35th International SAMPE Symposium, Anaheim CA (USA), Vol. 35, pp. 710–720, 1990.

- [28] Enders M. and Hopkins P., Developments in the fiber placement process. Proc. of 36th International SAMPE Symposium, San Diego, Vol. 36, pp. 778-790, 1991.
- [29] Martin J., Langone R., Pasanen M. and Mondo J., Cost-effective, automated equipment for advanced composite structure development and production. Tech. Rep. Automated Dynamics Corporation, 1997.
- [30] Evans D., Fiber placement. Tech. Rep. Cincinnati Machine, 2001.
- [31] Fuchiyama T. and Noda N., Analysis of thermal stress in a plate of functionally gradient material. JSAE Rev, 6, pp. 263-268, 1995.
- [32] Suresh S. and Mortensen A., Fundamentals of functionally graded materials. IOM Communications Limited, 1998.
- [33] Mahfuz H., Islam M.S., Rangari V.K., Saha M.C. and Jeelani S., Response of sandwich composites with nanophased cores under flexural loading. Composites, Part B, Vol. 35, pp. 543-50, 2004.
- [34] Sankar B.V., An elasticity solution for functionally graded beams. Comp Sci. Tech, Vol. 61, pp. 689-96, 2001.
- [35] Sankar B.V. and Tzeng J.T., Thermal stresses in functionally graded beams. AIAA J, Vol. 40, No. 6, pp. 1228-32, 2002.
- [36] Venkataraman S. and Sankar B.V., Analysis of sandwich beams with functionally graded core. Proc. of 42nd AIAA structures, structural dynamics and materials conference, Seattle, AIAA paper 2001-1281, 2001.
- [37] Apetre N.A., Sankar B.V. and Venkataraman S., Indentation of a sandwich beam with functionally graded core. Proc. of 43rd AIAA Structures, Structural Dynamics and Materials Conference. Denver CO (USA), AIAA paper 2002-1683, 2002.
- [38] Apetre N.A., Sankar B.V. and Ambur D.R., Low-velocity impact response of sandwich beams with functionally graded core. Int. J. Solids Struct, Vol. 43, pp. 2479-2496, 2006.
- [39] Apetre N.A., Sankar B.V. and Ambur D.R., Functionally-graded sandwich core with arbitrary variation in properties. Proc. of 44th AIAA Structures, Structural Dynamics and Materials Conference. Norfolk VA (USA), AIAA paper 2003-1947, 2003.
- [40] Zhu H and Sankar B.V., A combined Fourier series-Galerkin method for the analysis of functionally graded beams. J Appl Mech, Vol. 71, No. 3, pp. 421-424, 2004.
- [41] Reid R.G. and Paskaramoorthy R., An extension to classical lamination theory for use with functionally graded plates. Compos Struct, Vol. 93, pp. 639-648, 2011.
- [42] Icardi U. and Ferrero L., A new tailoring optimization approach for improving structural response and energy absorption capability of laminated and sandwich composites. J. Mech Mater Struct, Vol. 3, No. 4, pp. 729-60, 2008.
- [43] Icardi U. and Ferrero L., Laminated and sandwich panels subject to blast pulse loading. J. Mech. Mater. Struct., Vol. 4, No. 9, pp. 1573-94, 2009.
- [44] Icardi U. and Ferrero L., Optimisation of sandwich panels with functionally graded core and faces. Composites Science & technology, Vol. 69, pp. 575-585, 2009.
- [45] Icardi U. and Ferrero L., Optimisation of sandwich panels to blast pulse loading. J. sandwich Structures & Materials, Vol. 12, No. 5, pp. 521-550, 2010.
- [46] Icardi U., Higher-order zig-zag model for analysis of thick composite beams with inclusion of transverse normal stress and sublaminates approximations, Composites: Part B, Vol. 32, pp. 343-354, 2001.
- [47] Icardi U. and Sola F., Easy development of a zig-zag model with variable representation of displacements across the thickness, Submitted to J. of Comput. & Appl. Res. in Mech. Eng.
- [48] Cho M., Kim K.O. and Kim M.H., Efficient higher-order shell theory for laminated composites. Composite Structures, Vol. 34, pp. 197-212, 1996.
- [49] Li X.Y and Liu D., Generalized laminate theories based on double superposition hypothesis. Int J Numer Meth Eng, Vol. 40, pp. 1197-212, 1997.
- [50] Zhen W. and Wanji C., A global-local higher order theory for multilayered shells and the analysis of laminated cylindrical shell panels, Composite Structures, Vol. 84, pp. 350-361, 2008.
- [51] Hyer M. and Lee H., The use of curvilinear fiber format to improve buckling resistance of composite plates with central circular holes. Composite Structures, Vol. 18, pp. 239-261, 1991.
- [52] Icardi U., Extension of a new tailoring optimisation technique to sandwich shells with laminated faces. Struct. Eng. and Mech., in press
- [53] Gibson L.J. and Ashby M.F., Cellular solids. Pergamon, Oxford, 1988.
- [54] Pagano N.J., Exact solutions for rectangular bidirectional composites and sandwich plates. J Compos. Mater., Vol. 4, pp. 20-34, 1970.
- [55] Hou J.P., Petrinic N. and Ruiz C., A Delamination Criterion for Laminated Composites Under Low-Velocity Impact. *Composites Science and Technology*, Vol. 61, pp. 2069-2074, 2001.

STATIC AND LOW FREQUENCY EVALUATION OF A THIN-FILM CONTACT FORCE SENSOR

Andrea Manuello Bertetto

Silvia Meili

Department of Mechanical, Chemical and Materials Engineering, University of Cagliari
via Marengo 2 - 09123 Cagliari, Italy

ABSTRACT

Service robots meant for human interaction need to simulate touch sensitivity, or contact force is to be known for control purposes. Measurements have thus to be quick, reliable and have a good frequency response in the required band. Moreover, minimal dimensions are required, as well as an evenly responsive contact surface. We chose a commercial contact force sensor which exhibits all of these features: minimal overall dimensions, uniform measuring area, customizable range and certified linearity, no drift and hysteresis or good frequency response features. We present here our static and dynamic calibration results along with peculiarities we came across. We wanted to assess if the sensors would exhibit linear behaviour, no drift nor hysteresis as stated by the manufacturer. The sensors should offer uniform sensitivity over the sensing area, so matching section weights can be used to load them for static calibration. Besides this simple method, we also used hydrostatic pressure in a sealed chamber, getting rid of errors uncertainties coming from loading over the edges or from point loading. This method also allows for dynamic testing, as impulse loading or different frequency pressure waves can be applied to the sensors. Calibration curves will be presented in different loading conditions, focusing our attention on loading velocity and frequency. Although sensor gain exhibits remarkable changes after repeated loading, its smooth and continuous output make us think it can still be put in a feedback chain.

Keywords: piezoresistive sensors, calibration, pneumatics, control, mechatronics

1 INTRODUCTION

Tactile sensors are a key component in robotic applications requiring human-robot interaction. Surface force or pressure transducers are used both for simple qualitative touch sensing, thus requiring a simple binary output, and for quantitative continuous gauging of contact pressure through analog voltage signals.

Pressure sensors allow for feedback controlling the system or safety monitoring to prevent harm to the user or damage of the device. A large class of service robots are meant for biomedical applications or care purposes as artificial hands, gait orthoses or rehabilitation and wellness devices in general.

A few commercial, low cost force sensors are available on the market [1, 2] and for some of them the producer doesn't guarantee an acceptable repeatability, not to mention linear or hysteresis free response. Many authors have studied such sensors' behaviour [3-6]. Sensors come in many sizes and shapes, given their prior requirement not to be thick or intrusive, and they work based on a few different principles, being them capacitive or piezoresistive, being too bulky elastic load cells. As for the shape/size issue, thin sensors' dimensions span from a few millimetres, allowing for point contact measurements i.e. for artificial fingers or point-wise, spherical-ended tools, to large arrays of tenth of centimetres side for larger contact areas, i.e. for stress distribution under the plantar sole or over the dorsal area or between soil and foundations [7, 8] for civil engineering applications.

Our main target in considering these piezo-resistive sensors was a soft-robot application, i.e. a deformable pneumatic

Contact author: Silvia Meili

E-mail: s.meili@dimcm.unica.it

device intended to boost leg circulation of paraplegic patients by applying an intermittent massage and at the same time imposing passive ankle dorsiflexion [9]. Measuring the contact force between the actuator and the patients skin should give reliable and quick response in order to tune contact pressure. Thin-film sensors to be mounted under the actuator inner surface are eligible for this task, thus need to be tested and characterized.

The FlexiForce® sensors from Tekscan Inc., Boston, MA, US seemed to us a good commercial and affordable offer for measuring small forces, also given their short response time ($<5 \mu\text{s}$). Good frequency response in the low-frequency band of our interest ($0.5\div 1$ Hz, meaning the gait frequency) is required.

2 SENSORS AND TESTBENCH

2.1 SENSORS PROPERTIES

The Tekscan FlexiForce® sensor is made of two polyester sheets patterned with piezoresistive ink electrodes (see Figure 1).

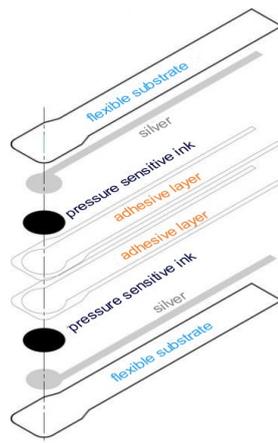


Figure 1 Sensor construction.

The sensor acts as a force sensing resistor in an electrical circuit. The lower the load, the higher the resistance. When a force is applied to the sensor, this resistance decreases.



Figure 2 The amplifier based on the suggested op-amp drive circuit.

As the manufacturer suggests, we used a custom op-amp drive circuit (see Figure 2) following the given scheme to produce a linear voltage output between $0\div 10$ volts and allowing for DC power supplying eight sensors at once, being in turn feeded at 5 V through a USB cable.

The A201 model we are studying has a 9.53 mm diameter sensing area, and a $0\div 4.4$ N force range. These piezoresistive sensors are guaranteed for a linear output, acceptable repeatability and low hysteresis (see Table I).

Table I – FlexiForce® A201 data sheet

Linearity (Error)	$< \pm 3\%$
Repeatability	$< \pm 2.5\%$ of Full Scale
Hysteresis	$< 4.5\%$ of Full Scale
Drift	$< 5\%$ per Logarithmic Time Scale
Response Time	$< 5\mu\text{s}$

2.1 STATIC CALIBRATION

Two methods were adopted: dead weights laying in the sensors surface, and hydrostatic air pressure in a sealed container. For applying loads on the sensors surface a plastic shim (Figure 3) was placed between the sensing area and the weight bottom when contact areas would not match, and to uniformly distribute loading when the contact surface was not smooth, thus avoiding kinks or stress concentrations. Of course such care cannot be taken in actual applications.



Figure 3 The shim used to distribute the load.

A pressure vessel (a sealed bottle, see Figure 5) was used to test the sensors using hydrostatic pressure, thus getting rid of all those issues tied to stress distribution, and should give as clean output values as possible.

2.1.1 Drift

Four increasing weights were laid over the sensor surface for 2 hours each to assess drift, and pressure values were continuously monitored and recorded by means of a NI cDAQ-9172 CompactDAQ chassis, equipped with a 9219 4-Channel, 24-Bit, analog input module. A sample of acquired data is shown in figure 5. Loading conditions changes are not associated to the gap in output voltage values.

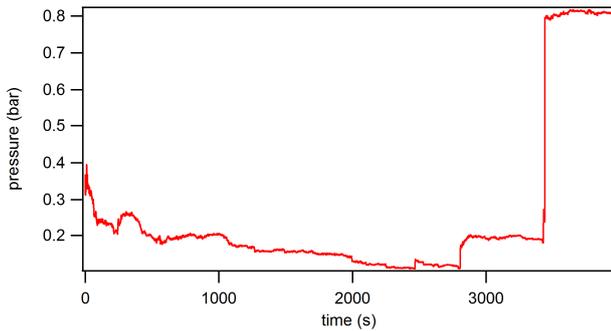


Figure 4 An example of the drift trend in about one hour time.

2.1.2 Hydrostatic pressure loading

Static loading curves, some of them followed by unloading, were all recorded at the same velocity, about the same as the one adopted in working conditions, and after the conditioning procedure suggested by the manufacturer: each sensor was repeatedly loaded and unloaded at about 110% full scale, and tested almost immediately not to lose conditioning effects.

It is to be noted how each curve is smooth and how they are perfectly scaled to each other although they do not match.



Figure 5 The sealed plastic bottle used for hydrostatic pressure testing.

Nevertheless, it doesn't seem that repeatability can be achieved anyway in spite of not changing the testing conditions.

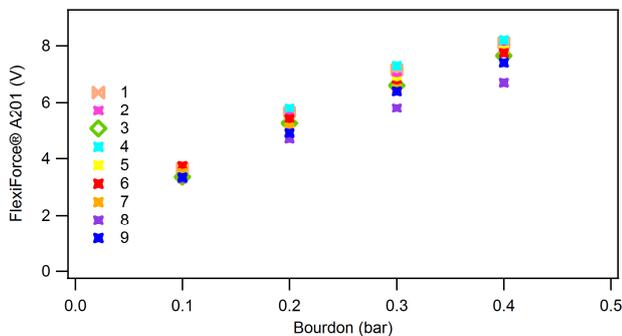


Figure 6 Static-loading curves (sensor voltage output vs high-precision Bourdon gauge): the gain is not related to the loading sequence.

Moreover, the test sequence is not related to the gain increasing or decreasing.

As for linearity, in Figure 7 are shown two rectified calibration curves and their coefficients. The power law $p=(V/12)^2$ seems to give better fitting here.

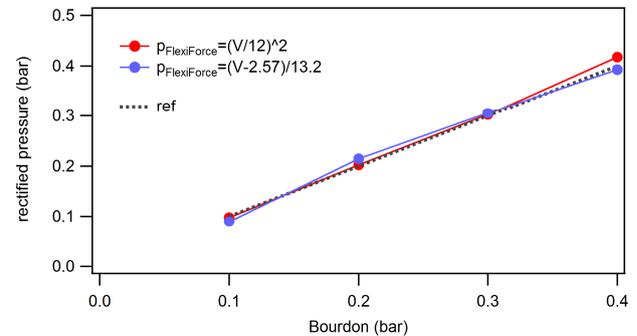


Figure 7 Calibration curves for one piezoresistive sensor.

2.2 DYNAMIC CALIBRATION

2.2.1 Purpose

We wanted to assess the loss of calibration occurring over multiple loading cycles, if any, as static loading and drift curves suggest. This would minimize the sensor utility for repeated, long-term load measurements not to speak of feedback controlling the system.

A preliminary loading-unloading curve is shown in Figure 5. During the test the sensor behaves in a satisfactory and reliable way, without any appreciable response changes.

2.2.2 Sine wave loading

The sensor was cycle-loaded at low frequencies by means of one of the two analog outputs (16-bit, 833 kS/s) of a NI-6221 USB module piloting at 0÷10 V DC a Norgren VP10 E/P converter with an air consumption of 0.85 NI/min and a response time less than 0.5 s for 10-90% step change. Experimental setup is shown in Figure 8.

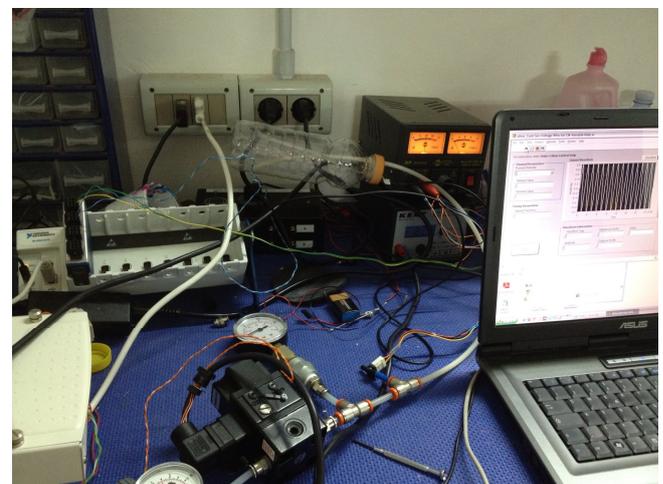


Figure 8 Experimental setup

The proportional valve was power supplied at 24 V DC. A few loading conditions were chosen to assess the sensors performance: pressure was imposed to oscillate around a mean value of 0.5 bar at frequencies of 0.5 Hz and 1 Hz, and 0.2 bar amplitude (0.4±0.6 bar). Air pressure was also monitored upstream of the container by means of a 0-15 psi Honeywell 24 PC Series piezo-resistive differential pressure gauge with 15 mV/psi sensitivity.

The green curves in Figure 9a, 9b below give the line pressure measured by the transducer, while the blue ones are the output of the thin-film sensor inside the pressure bottle. They do not match in amplitude nor in global trend, generally overestimating the measured pressure.

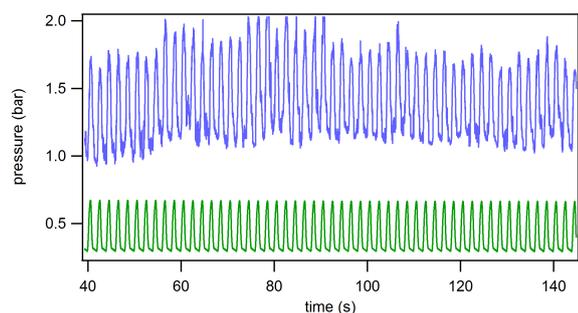
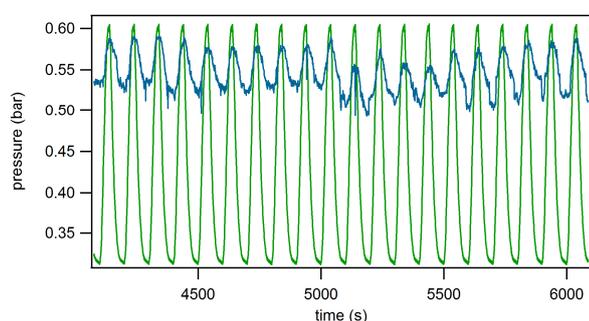


Figure 9 Power function calibration coefficients for a loading-unloading curve. A cyclic contact measurement (in blue) is also shown for comparison.

2.3 CONTACT TESTING

In Figure 9 below a power function calibration is given for a loading-unloading curve in the 0±0.4 bar range.

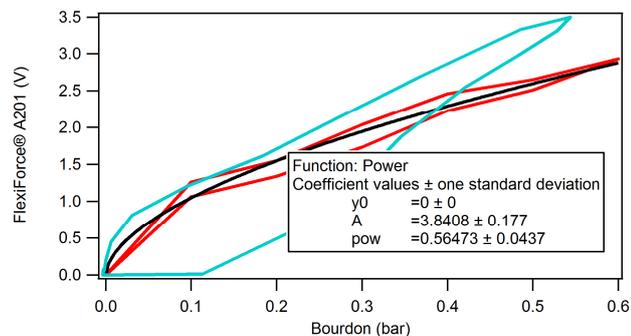


Figure 10 Power function calibration coefficients for a loading-unloading curve. A cyclic contact measurement (in blue) is also shown for comparison.

The blue curve was obtained interposing the thin-film sensor between a plastic pipe and the inflatable actuator (a cuff-bladder system) of the device in [5]. In Figure 11 a cyclic loading-unloading curve was obtained by inflating-deflating the cuff. The black curve comes from testing the system on a arm. The red curve exhibits higher hysteresis.

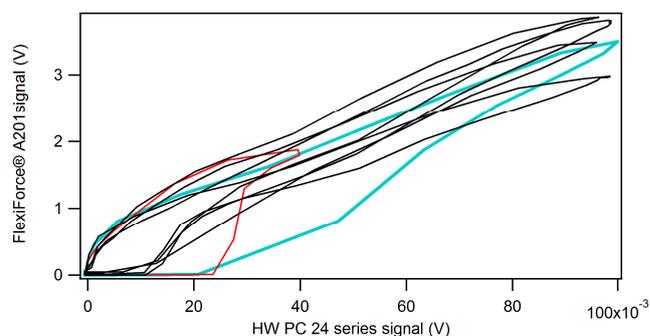


Figure 11 Qualitative trend of the measured pressure contact between a plastic tube and the cuff-bladder system in [9] (blue curve) and cyclic loading-unloading curves from contact pressure measurements between the same actuator as above and a human arm. Air pressure is given by the analog output of the Honeywell differential transducer.

3 RESULTS

The drift curves doesn't seem to be satisfactory, as they present sudden gaps that cannot be foreseen or avoided, or simply don't give repeatable measures being measured pressure unchanged.

The static loading curves as well exhibit no repeatability nor linearity. A question could be risen about the opportunity to use a linear or a power function to give the best fit. Both solutions are presented. The pressure range could lead towards one ore the other.

As for low-frequency cyclic loading, the sensors' behaviour doesn't seem to be reliable either, as unexpected and unjustified raises and drops of the output occur, still being it smooth.

4 CONCLUSIONS

These observations make us think that the studied thin-film piezo-resistive sensors deserve a further systematic investigation, being their advantages in terms of size, smooth output, apparently good frequency response -in the band of use and investigation- and sensitivity still predominant on their poor repeatability, which could be rectified i.e. by an on-the-fly tuning.

5 ACKNOWLEDGMENTS

This work was supported by a grant from Italian MIUR (Italian Ministry of Education, University and Research, PRIN 2008)

REFERENCES

- [1] Vecchi F., Freschi C., Micera S., Sabatini A. M., Dario P. and Sacchetti R., Experimental Evaluation of Two Commercial Force Sensors for Applications in Biomechanics and Motor Control. *5th Annual Conference of the International Functional Electrical Stimulation Society*, Aalborg, DK, pp. 4-7, 2000.
- [2] Hollinger A. and Wanderley M.M., Evaluation of Commercial Force-Sensing Resistors. *International conference on new interfaces for musical expression NIME06*, Paris, 2006.
- [3] Castro M. C. F. and Cliquet Jr, A., A low-cost instrumented glove for monitoring forces during object manipulation. *IEEE Transactions on Robotics*, Vol. 5, No. 2, pp. 140-147, 1997.
- [4] Vidal-Verdú F., Barquero M. J., Castellanos-Ramos J., Navas-González R., Sánchez J. A., Seron J. and Garcia-Serzo A. A Large Area Tactile Sensor Patch Based on Commercial Force Sensors. *Sensors*, Vol. 11, No. 5, pp. 5489–5507, 2011.
- [5] Freschi C., Vecchi F., Micera S., Sabatini A. M., Dario P. and Sacchetti R., Force Control during Grasp using FES Techniques: Preliminary Results. *5th Annual Conference of the International Functional Electrical Stimulation Society*, Aalborg, DK, pp. 6-9, 2000.
- [6] Otto J.K., Brown, T.D., and Callaghan J.J, Static and Dynamic Response of a Multiplexed-array Piezoresistive Contact Sensor, *Experimental Mechanics*, Vol. 39, No. 4, December 1999.
- [7] Shaw A. J., Davis B. A., Collins M. J., and Carney L. G.. A Technique to Measure Eyelid Pressure Using Piezoresistive Sensors. *IEEE Transactions On Biomedical Engineering*, Vol. 56, No. 10, October 2009.
- [8] Palmer M. C., O'Rourke T. D., Olson N. A., Abdoun T., Ha D., Rourke M. J. O. Tactile Pressure Sensors for Soil-Structure Interaction Assessment. *Journal of Geotechnical and Geoenvironmental Engineering*, Vol. 135, No. 11, November 1, 2009.
- [9] Manuello Bertetto A., Meili S., Concu A. and Crisafulli A., An Inflatable Pneumatic System for Blood Pressure Recovery, *Mechanics Based Design of Structures and Machines*, Vol. 40, No. 4, pp. 506-518, 2012.

TEMPLATE FOR PREPARING PAPERS FOR PUBLISHING IN INTERNATIONAL JOURNAL OF MECHANICS AND CONTROL

Author1* Author2**

* affiliation Author1

** affiliation Author2

ABSTRACT

This is a brief guide to prepare papers in a better style for publishing in International Journal of Mechanics and Control (JoMaC). It gives details of the preferred style in a template format to ease paper presentation. The abstract must be able to indicate the principal authors' contribution to the argument containing the chosen method and the obtained results. (max 200 words)

Keywords: keywords list (max 5 words)

1 TITLE OF SECTION (E.G. INTRODUCTION)

This sample article is to show you how to prepare papers in a standard style for publishing in International Journal of Mechanics and Control.

It offers you a template for paper layout, and describes points you should notice before you submit your papers.

2 PREPARATION OF PAPERS

2.1 SUBMISSION OF PAPERS

The papers should be submitted in the form of an electronic document, either in Microsoft Word format (Word'97 version or earlier).

In addition to the electronic version a hardcopy of the complete paper including diagrams with annotations must be supplied. The final format of the papers will be A4 page size with a two column layout. The text will be Times New Roman font size 10.

2.2 DETAILS OF PAPER LAYOUT

2.2.1 Style of Writing

The language is English and with UK/European spelling. The papers should be written in the third person. Related work conducted elsewhere may be criticised but not the individuals conducting the work. The paper should be comprehensible both to specialists in the appropriate field and to those with a general understanding of the subject. Company names or advertising, direct or indirect, is not permitted and product names will only be included at the discretion of the editor. Abbreviations should be spelt out in full the first time they appear and their abbreviated form included in brackets immediately after. Words used in a special context should appear in inverted single quotation mark the first time they appear. Papers are accepted also on the basis that they may be edited for style and language.

2.2.2 Paper length

Paper length is free, but should normally not exceed 10000 words and twenty illustrations.

2.2.3 Diagrams and figures

Figures and Tables will either be entered in one column or two columns and should be 80 mm or 160 mm wide respectively. A minimum line width of 1 point is required at actual size. Captions and annotations should be in 10 point with the first letter only capitalised *at actual size* (see Figure 1 and Table VII).

Contact author: author1¹, author2²

¹Address of author1.

²Address of author2 if different from author1's address.

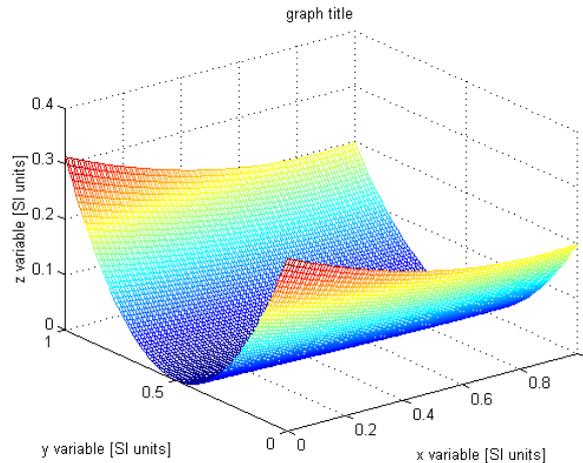


Figure 1 Simple chart.

Table VII - Experimental values

Robot Arm Velocity (rad/s)	Motor Torque (Nm)
0.123 10.123	
1.456 20.234	
2.789 30.345	
3.012 40.456	

2.2.4 Photographs and illustrations

Authors could wish to publish in full colour photographs and illustrations. Photographs and illustrations should be included in the electronic document and a copy of their original sent. Illustrations in full colour ...

2.2.5 Equations

Each equation should occur on a new line with uniform spacing from adjacent text as indicated in this template. The equations, where they are referred to in the text, should be numbered sequentially and their identifier enclosed in parenthesis, right justified. The symbols, where referred to in the text, should be italicised.

- point 1
 - point 2
 - point 3
- 1. numbered point 1
- 2. numbered point 2
- 3. numbered point 3

$$W(d) = G(A_0, \sigma, d) = \frac{1}{T} \int_0^{+\infty} A_0 \cdot e^{-\frac{d^2}{2\sigma^2}} dt \quad (1)$$

3 COPYRIGHT

Authors will be asked to sign a copyright transfer form prior to JoMaC publishing of their paper. Reproduction of any part of the publication is not allowed elsewhere without permission from JoMaC whose prior publication must be cited. The understanding is that they have been neither previously published nor submitted concurrently to any other publisher.

4 PEER REVIEW

Papers for publication in Jo MaC will first undergo review by anonymous, impartial specialists in the appropriate field. Based on the comments of the referees the Editor will decide on acceptance, revision or rejection. The authors will be provided with copies of the reviewers' remarks to aid in revision and improvement where appropriate.

5 REFERENCES (DESCRIPTION)

The papers in the reference list must be cited in the text. In the text the citation should appear in square brackets [], as in, for example, "the red fox has been shown to jump the black cat [3] but not when...". In the Reference list the font should be Times New Roman with 10 point size. Author's first names should be terminated by a 'full stop'. The reference number should be enclosed in brackets.

The book titles should be in *italics*, followed by a 'full stop'. Proceedings or journal titles should be in *italics*. For instance:

REFERENCES (EXAMPLE)

- [1] Smith J., Jones A.B. and Brown J., *The title of the book*. 1st edition, Publisher, 2001.
- [2] Smith J., Jones A.B. and Brown J., The title of the paper. *Proc. of Conference Name*, where it took place, Vol. 1, paper number, pp. 1-11, 2001.
- [3] Smith J., Jones A.B. and Brown J., The title of the paper. *Journal Name*, Vol. 1, No. 1, pp. 1-11, 2001.
- [4] Smith J., Jones A.B. and Brown J., *Patent title*, U.S. Patent number, 2001.

International Journal of Mechanics and Control – JoMaC
Published by Levrotto&Bella
TRANSFER OF COPYRIGHT AGREEMENT

<p>NOTE: Authors/copyright holders are asked to complete this form signing section A, B or C and mail it to the editor office with the manuscript or as soon afterwards as possible.</p>	<p><i>Editor's office address:</i> Andrea Manuello Bertetto Elvio Bonisoli <i>Dept. of Mechanics</i> <i>Technical University – Politecnico di Torino</i> <i>C.so Duca degli Abruzzi, 24 – 10129 Torino – Italy</i> <i>e_mail: jomac@polito.it</i> <i>fax n.: +39.011.564.6999</i></p>
--	---

The article title:

By: _____

To be Published in *International Journal of Mechanics and Control JoMaC*
Official legal Turin court registration Number 5320 (5 May 2000) - reg. Tribunale di Torino N. 5390 del 5 maggio 2000

- A Copyright to the above article is hereby transferred to the JoMaC, effective upon acceptance for publication. However the following rights are reserved by the author(s)/copyright holder(s):
1. All proprietary rights other than copyright, such as patent rights;
 2. The right to use, free or charge, all or part of this article in future works of their own, such as books and lectures;
 3. The right to reproduce the article for their own purposes provided the copies are not offered for sale.
- To be signed below by all authors or, if signed by only one author on behalf of all co-authors, the statement A2 below must be signed.*

A1. All authors:

SIGNATURE _____ DATE _____ SIGNATURE _____ DATE _____

PRINTED NAME _____ PRINTED NAME _____

SIGNATURE _____ DATE _____ SIGNATURE _____ DATE _____

PRINTED NAME _____ PRINTED NAME _____

A2. One author on behalf of all co-authors:

"I represent and warrant that I am authorised to execute this transfer of copyright on behalf of all the authors of the article referred to above"

PRINTED NAME _____

SIGNATURE _____ TITLE _____ DATE _____

B. The above article was written as part of duties as an employee or otherwise as a work made for hire. As an authorised representative of the employer or other proprietor. I hereby transfer copyright to the above article to *International Journal of Mechanics and Control* effective upon publication. However, the following rights are reserved:

1. All proprietary rights other than copyright, such as patent rights;
2. The right to use, free or charge, all or part of this article in future works of their own, such as books and lectures;
3. The right to reproduce the article for their own purposes provided the copies are not offered for sale.

PRINTED NAME _____

SIGNATURE _____ TITLE _____ DATE _____

C. I certify that the above article has been written in the course of employment by the United States Government so that no copyright exists, or by the United Kingdom Government (Crown Copyright), thus there is no transfer of copyright.

PRINTED NAME _____

SIGNATURE _____ TITLE _____ DATE _____

CONTENTS

- 3 Modelling and Temperature Control of Shape Memory Alloys with Fast Electrical Heating**
R. Velázquez and E.E. Pissaloux
- 11 Vehicles Headlamps Glare Effect Removal**
G. Ligios and A. Manuello Bertetto
- 37 A Dry Friction Model and Robust Computational Algorithm for Reversible or Irreversible Motion Transmissions**
L. Borello and M.D.L. Dalla Vedova
- 49 Design Procedure of a Breech Mechanism with Double Helix Shaft**
C. Delprete, E. Bonisoli and D. Tomasi
- 57 Modular Reconfigurable Anthropomorphic Gripper with Three Fingers for Robots-Design, Prototype and Simulation**
C.C. Moldovan and I. Staretu
- 63 Velomobiles: Design Guidelines**
P. Baldissera, C. Delprete and M. Tirelli
- 73 A Methodology to Design Robotic Arms for Service Tasks since Early Design Stage**
C. Castejón, G. Carbone, J.C. García-Prada and M. Ceccarelli
- 85 Vehicle Parameters Identification and States Estimation**
D. Capra, M. Velardocchia and A. Vigliani
- 91 Application of a New Tailoring Optimization Technique to Laminated and Sandwich Plates and to Sandwich Spherical Panels**
U. Icardi and F. Sola
- 107 Static and Low Frequency Evaluation of a Thin-Film Contact Force Sensor**
A. Manuello Bertetto and S. Meili