

Quaderni del **DIEM – GMA**

Atti di giornate di studio – 2

A cura di:

U. Meneghetti, A. Maggiore, V. Parenti Castelli

Coordinatore di redazione:

Alessandro Zanarini

DIEM

Dipartimento di Ingegneria delle
Costruzioni Meccaniche, Nucleari, Aeronautiche e di
Metallurgia
www.diem.unibo.it

GMA

Gruppo di Meccanica Applicata
http://wpage.unina.it/dellaval/GMA/GMA_home.htm

Seconda giornata di studio Ettore Funaioli

18 luglio 2008

*A cura di:
Umberto Meneghetti, Alberto Maggiore e Vincenzo Parenti Castelli*

ASTE RISCO
ASTE RISCO
ASTE RISCO
EDIZIONI



*Proprietà letteraria riservata.
© Copyright 2009 degli autori
Tutti i diritti riservati*

*Seconda giornata di studio Ettore Funaioli - 18 luglio 2008 / A cura di
Umberto Meneghetti, Alberto Maggiore e Vincenzo Parenti Castelli
Bologna: 2009 - p. 346; 17 cm.*

ISBN 978-88-86909-53-2

Versione elettronica disponibile alla pagina
<http://amsacta.cib.unibo.it/archive/00002552/>

Stampa a richiesta eseguita da:
Asterisco Snc Tipografia Digitale
Via Belle Arti, 31 a/b – 40126 Bologna
Tel 051 236866 – Fax 051 261105
mail: grafica@asteriscosnc.it
www.asteriscosnc.it

Indice

Prefazione

- A. Martini, M. Troncossi, M. Carricato, A. Rivola* 1
On the elastodynamic behaviour of balanced closed-loop mechanisms:
numerical analysis of a four-bar linkage
- E. Mucchi, G. Dalpiaz* 13
Simulation of gear pump dynamic behaviour by an integrated LP/FE
model
- A. Zanarini* 31
Full-field ESPI techniques in structural dynamics
- E. Dragoni, D. Castagnetti, A. Spaggiari* 57
Analisi a collasso di giunzioni incollate: modellazione efficiente agli
elementi finiti e convalida sperimentale
- A. Freddi, G. Olmi* 67
Progetto a fatica di ruote dentate ad elevate prestazioni
- G. Minak, D. Ghelli, A. Zucchelli* 93
Sulla resistenza residua a torsione dopo impatto di tubi in materiale
composito
- A. Strozzi, A. Baldini, M. Giacomini, R. Rosi* 103
Three examples of inadequacy of the Williams asymptotic method in
predicting the singularity strength in elastic problems
- G. Medri* 113
Design drivers per lo sviluppo di nuovo prodotto
- G. Vassura, A. Macchelli* 121
Project-based co-tutored mechatronic learning
- A. Zucchelli, G. Minak, D. Ghelli* 133
Studio numerico e sperimentale del comportamento meccanico di lastre
in acciaio ed in acciaio porcellanato soggette ad impatto a bassa velocità
- A. O. Andrisano, G. Barbanti, M. Pellicciari, F. Leali, M. Faretra,
F. Gherardini, F. Pini, A. Vergnano* 153
Metodi di progettazione integrata e simulazione di sottogruppi di
macchine alternative
- L. Donati, L. Tomesani, M. Schikorra, A. E. Tekkaya* 165
Microstructure prediction in AA6060 hot profile extrusion

<i>G. Tani, C. Mantega</i>	181
Misura sperimentale e previsione delle forze di taglio, per materiali dalle caratteristiche meccaniche variabili	
<i>L. Orazi, A. Fortunato, G. Tani, G. Cuccolini, A. Ascari, G. Campana</i>	193
Modellazione del processo di trattamento termico superficiale mediante fascio laser	
<i>R. Squatrito, I. Todaro, A. Morri, L. Tomesani</i>	207
Simulazioni di processi di colata di componenti industriali in alluminio di alta complessità	
<i>R. Di Gregorio</i>	221
Review and comparison of the metrics proposed for measuring the distance between two rigid-body poses	
<i>E. Dragoni, G. Scirè Mammano, I. Spinella</i>	235
Dimensionamento meccanico, controllo e prototipazione di attuatori a memoria di forma	
<i>C. Innocenti</i>	261
Una dimostrazione didatticamente efficace del teorema di Eulero-Savary	
<i>R. Vertechy, G. Berselli, V. Parenti Castelli, G. Vassura</i>	271
A new procedure for the optimization of a dielectric elastomer actuator	
<i>A. Benedetti, P. G. Molari, P. Morelli</i>	283
Sull'innesco della frattura in contatti non Hertziani	
<i>S. Delvecchio, G. D'Elia, G. Dalpiaz</i>	293
Application of advanced vibration processing techniques in i. c. engine cold tests	
<i>U. Meneghetti, A. Maggiore</i>	313
Applicazione dell'oscillatore di Duffing per il monitoraggio delle macchine	
<i>M. Cocconcelli, R. Rubini</i>	327
Introduction to a simply and fast algorithm for variable speed bearing diagnostics	
<i>E. Prati, A. Tasora, M. Silvestri</i>	337
Collaborazioni fra l'Università degli Studi di Parma e centri di ricerca internazionali nell'ambito della meccanica applicata	
Indice degli autori	345

PREFAZIONE

“Non c’è modo migliore per onorare i nostri Maestri che continuarne l’opera e dedicare alla loro memoria i risultati della nostra attività”: così ha efficacemente commentato l’iniziativa delle “Giornate Ettore Funaioli” il Prof. Pier Paolo Diotallevi, Preside della Facoltà di Ingegneria di Bologna, portando il saluto della Facoltà ai partecipanti.

Con questo spirito, il 18 luglio 2007 si svolse la "Prima giornata di studio in onore del Prof. Ettore Funaioli", alla quale furono invitati i Colleghi che ebbero occasione di collaborare, direttamente o indirettamente, con Ettore Funaioli, per tanti anni Professore Ordinario di Meccanica applicata alle macchine nella Facoltà di Ingegneria di Bologna.

Dato l’interesse suscitato, nel 2008 l’iniziativa è stata ripetuta, estendendo l’invito a partecipare – con riferimento al nascente "Coordinamento della Meccanica" – ai Ricercatori di Meccanica dell’Emilia Romagna e, in particolare, a coloro che operano nei settori della Meccanica applicata alle macchine, della Costruzione di macchine, del Disegno di macchine e della Tecnologia meccanica.

Allo scopo di favorire l’incontro e la reciproca conoscenza fra i Ricercatori dei settori affini più sopra menzionati, la scelta del contenuto delle memorie è stata lasciata al criterio dei presentatori: il presente volume raccoglie, pertanto, sia memorie scientifiche vere e proprie, sia illustrazioni di attività di Gruppi di ricerca, sia esposizioni dello stato di avanzamento di ricerche in corso.

Un sentito ringraziamento porgiamo al Prof. Pier Paolo Diotallevi, Preside della Facoltà di Ingegneria, e al Prof. Emilio Ferrari, Direttore del DIEM, per la fattiva collaborazione e per le parole con cui hanno voluto inaugurare la “Seconda Giornata di studio Ettore Funaioli”; un caloroso ringraziamento porgiamo poi agli Autori delle memorie qui raccolte, ai quali si deve per intero il successo della manifestazione.

Bologna, 31 gennaio 2009

ON THE ELASTODYNAMIC BEHAVIOUR OF BALANCED CLOSED-LOOP MECHANISMS: NUMERICAL ANALYSIS OF A FOUR-BAR LINKAGE

Alberto Martini

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: alberto.martini6@unibo.it

Marco Troncossi

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: marco.troncossi@unibo.it

Marco Carricato

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: marco.carricato@mail.ing.unibo.it

Alessandro Rivola

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: alessandro.rivola@unibo.it

Abstract. *The purpose of this research is the analysis of the effects produced by different techniques of static balancing, e.g. mass redistribution and elastic compensation, on the dynamic behaviour of closed-loop mechanisms. In particular, this contribution focuses on the study of a planar four-bar linkage – considered as the simplest example of closed-loop mechanism. For comparison purposes, three models are studied: an unbalanced, a mass-balanced and an elastically-compensated linkage.*

The elastodynamic analysis is performed via a commercial multibody software. In particular, the flexibility of the links is modelled by the finite element method, whereas lumped parameters are employed to describe the compliance of the mechanical driving transmission.

The modal analysis of the three models is carried out, the natural frequencies and the mode shapes being computed for several positions of the mechanisms. Furthermore, the kineto-elastodynamic simulation is performed, the driving torque and the constraint reactions being derived for different input velocities.

The effects of the different balancing techniques are investigated and discussed. The performed simulations provide some interesting results, which justify the extension of the research to more complex closed-loop mechanisms.

Keywords: *Static balancing, Four-bar linkage, Elastodynamic modelling, Modal analysis.*

1. INTRODUCTION

A mechanism is said to be *statically balanced* (or *gravity compensated*) if zero external actions are required to maintain the mechanism at rest in all assumable configurations. This occurs when its total potential energy is constant. In such an instance, the actuators impart accelerations to the moving members, but they do not contribute to supporting their weight. A mechanism is said to be *dynamically balanced* if both the resultant vector and the resultant moment about an arbitrary pole of the inertia forces (addressed as *shaking force* and *shaking moment*, respectively) are equal to zero. In this case, in any operating condition only the gravity load of the moving links is transmitted to the fixed frame. Keeping the mechanism centre of mass (G) stationary throughout the motion is a sufficient condition to equal the shaking force to zero, and also a necessary requirement to make the shaking moment vanish. Since complete dynamic balancing is difficult to achieve in practice – the introduction of rotating inertias being generally required – a common strategy consists in compensating the shaking force only: in this case, the mechanism is said to be simply *force balanced*.

The main benefits deriving from mechanism balancing, depending on the specific technique adopted, may be summarized as follows [1]:

- the motor driving actions are expected to be lower and more uniform, due to the reduction of the force/torque peaks that are generally caused by unbalanced gravity and/or inertia forces: this leads to enhanced energy efficiency and allows for the choice of less powerful actuators, resulting in a reduction of the weight of the mechanism;
- fail-safe brakes, which must be installed in order to keep an unbalanced mechanism steady when the motors are turned off, can be reduced in size, thus obtaining further weight decrement and energy efficiency enhancement;
- the forces transmitted to the fixed frame decrease, thus leading to vibration reduction, decrement of joint wear and global higher durability.

In this paper, attention is focused on two techniques which are commonly employed in order to obtain gravity compensation, that is to say *mass balancing* and *elastic compensation*. According to the former strategy, the masses of the mechanisms are properly distributed so that G remains stationary, thus ensuring that the total potential energy, coinciding with the gravity potential energy, remains constant. The practical implementation of such a technique generally requires attaching counterweights to some links. Following the latter strategy, instead, elastic elements – generally linear extension springs – are introduced in the system so that the total potential energy, in this case given by the sum of gravity and elastic contributions, remains constant.

Two main advantages characterize mass balancing. Since this technique requires G to rest stationary, the shaking force vanishes, leading to a force-balanced mechanism. Furthermore, gravity compensation is achieved for any orientation (and magnitude) of the gravity vector with respect to the fixed frame, so that the balanced mechanism may be arbitrarily oriented. However, some drawbacks are generally brought about:

- due to the additional masses, the total weight and bulkiness of the system increase, causing a potential worsening of the mechanism functionality;
- the inertia actions associated with the counterweights increase the power required to actuators: hence, the size of the motors should be incremented as well, limiting one of the most important benefits given by gravity compensation;

- in the absence of driving or braking forces at the rest position, a variation of the payload relative to the designed value makes the mechanism uncontrollably move.

Static balancing by means of elastic elements does not lead to force-balanced mechanisms and gravity compensation is achieved for only one direction and magnitude of the gravity vector. However, the following benefits may be emphasized:

- the introduction of elastic elements generally involve only a small increment in the total weight and inertia of the system;
- thanks to the little amount of mass added to the mechanism, the driving actions are expected to decrease (although the joint reaction forces could be higher);
- a moderate variation of the payload in the absence of any driving action normally causes the system to reach a new equilibrium position close to the previous one; hence, the mechanism does not fall under the action of gravity forces and safety brakes are not strictly required.

Static balancing of *open-loop* mechanisms, such as serial manipulators, is a commonly employed practice: in fact, since actuators and brakes are generally located on the moving links, the application of a strategy which allows limiting their size represents an essential advantage [1]. As far as *closed-loop* linkages are concerned, e.g. parallel manipulators, several studies have been carried out. Actually, this kind of machines normally operates at significant velocities and accelerations, and gravity loads often result negligible with respect to inertia forces. As a consequence, static balancing may not seem motivated. Indeed, the benefits achieved by gravity compensation are controversial [2–4]. Nonetheless, static balancing may still be advantageous any time the output link has to sustain a remarkable weight at rest for significant time intervals, as the relevant reduction of static motor torques produces important energetic gains, increased security and benefits in terms of brake sizing. When the mechanism is employed in applications characterized by modest dynamics (with maximum accelerations one order of magnitude smaller than gravity), static balancing may also lead to less powerful actuators, gravity compensation positively influencing dynamic torques as well [5].

As for the elastodynamic aspects related to statically balanced *open-loop* mechanisms, they represent an argument of moderate interest, since these mechanisms are normally characterized by modest-dynamics operating conditions. Probably for this reason, only two studies concerning this problem may be found in literature, at least to the authors' knowledge. In [6] the elastodynamic effects of mass balancing on a commercial manipulator are studied, whereas in [7] an explicit comparison between mass balancing and elastic compensation on an industrial robot modelled via flexible parts is presented. Elastodynamic aspects concerning gravity-compensated *closed-loop* mechanisms should be taken into account with greater attention, due to the significant dynamics generally involved in the functioning of this kind of machines. Nonetheless, to the authors' knowledge, such aspects appear to have been neglected or not exhaustively studied up to now. Moreover, several assertions are usually assumed as true, even though no rigorous demonstrations may be found in the literature. For example, according to common assumptions, counterweights are expected to cause a lowering of the natural frequencies of the system, whereas the introduction of elastic elements is expected to cause a stiffening effect in the mechanism. Nonetheless only studies related to the elastodynamic behaviour of mass-balanced four-bar linkages are available, by either experimental analysis [3, 4] or theoretical modelling [8, 9] or both theoretical and experimental investigation [10–12]. To

the author's knowledge, no specific study concerning the elastodynamic behaviour of elastically-compensated mechanisms appears in the literature.

This research aims at carrying out a deeper investigation of the effects of mass balancing and elastic compensation on the elastodynamic behaviour of closed-loop mechanisms, with the final objective of determining an optimal balancing strategy. The contribution presented here, in particular, is meant to be a preliminary study, aimed at defining the most suitable analysis procedure and at identifying the most interesting areas of investigation. The paper focuses on the four-bar linkage, intended to be the simplest example of closed-loop mechanism. Attention is paid to the comparison between the elastodynamic behaviour of an unbalanced four-bar linkage, referred to as *4B-U*, and two gravity-compensated versions of the same mechanism obtained by mass balancing and elastic compensation, referred to as *4B-M* and *4B-E* respectively (Fig. 1). Modal and kineto-elastodynamic analyses are performed on the numerical models by means of commercial software. The preliminary results are shown and discussed.

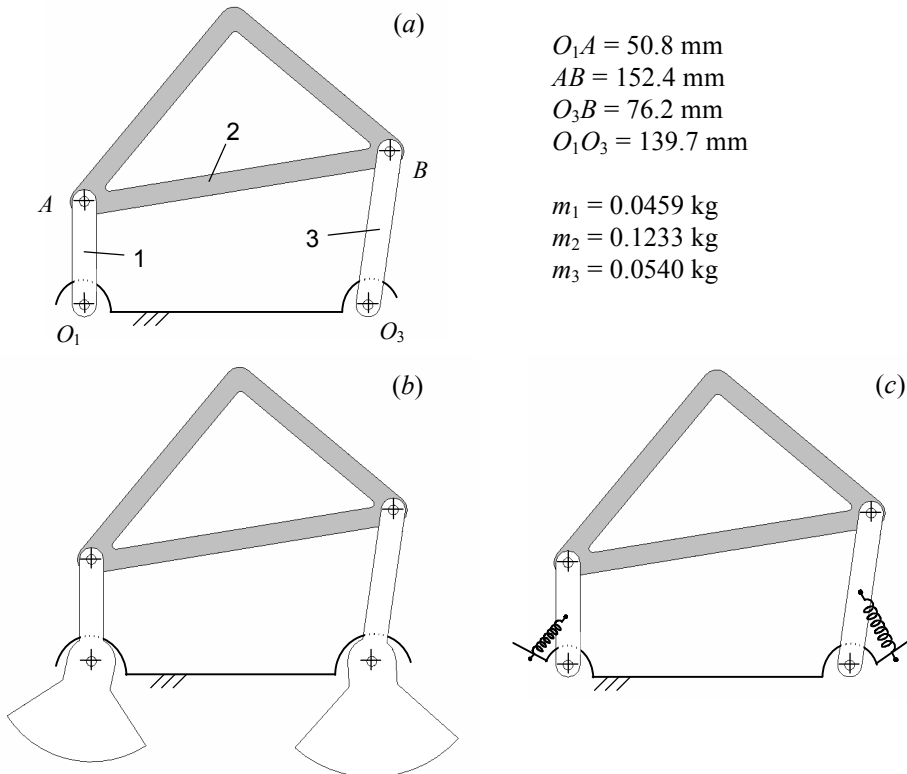


Figure 1. Schematics of the simulated four-bar linkages: (a) *4B-U*; (b) *4B-M*; (c) *4B-E*.

2. MATERIALS AND METHODS

The mechanism that is the object of this study is the crank-rocker four-bar linkage presented in [13]: the geometry and the mass properties of that linkage are used here to build up the *4B-U*. The *4B-M* is obtained by appending two counterweights to the crank and the rocker respectively, whereas in the *4B-E* balancing is achieved by means of two base-attached linear springs connected to the crank and the rocker. The “balancing parameters” of the *4B-M* may be determined following the Method of Linearly Independent Vectors as described in [13], whereas those of the *4B-E* may be obtained according to the approach presented in [14].

In order to perform the elastodynamic analysis, the flexibility of both the moving links and the mechanical driving transmission is considered. To this purpose, finite element models (FEM) of the crank, the rocker and the coupler are built via a commercial software (MD Patran/Nastran package) by using beam elements. The FEM are developed in such a way that each flexible part exhibits a planar behaviour, that is to say all its mode shapes lie on the plane of the mechanism. The models of links are then assembled within a multibody software (MD Adams), in order to form the final mechanism. At this stage, a lumped stiffness parameter is added to model the driving transmission. Further mass and stiffness lumped parameters are added to both the *4B-M* and the *4B-E* in order to model the balancing elements.

Modal analyses and kineto-elastodynamic (KED) simulations of the three four-bar linkages are performed by MD Adams. In particular, the natural frequencies and the mode shapes of the mechanisms are computed and investigated. Since these modal parameters change along the mechanism position, each linkage is driven through different configurations: 72 different positions are considered by using a crank angular increment of 5 degrees. The eigenvalues and eigenvectors are computed considering a static condition (i.e. nought velocity and acceleration are imposed on the driving shaft) for all the configurations. Comparisons of the modal parameters are made both between the three mechanisms and between different configurations of the same mechanism: in particular, mode shapes are analysed by visual inspection and by calculating the Modal Assurance Criterion (MAC). As far as the KED simulations are concerned, damping is introduced in the model: a constant damping ratio is chosen for all flexible links, whereas a proportional damping is considered for lumped stiffness parameters. A velocity law is imposed on the driving shaft, making the crank rotate at a constant angular velocity after a run-up revolution from zero to the full speed. Different working conditions are tested, corresponding to three values of the input velocity: 60, 300 and 600 RPM (i.e. 1, 5 and 10 Hz). The torque acting on the driving shaft and the joint reaction forces are computed and analysed.

3. RESULTS

Figure 2 shows the first natural frequencies of the three mechanisms as a function of the crank angular position. The mean values of the natural frequencies of the three four-bar linkages (averaged over the analysed positions) are shown in Table 1. The main remark is that, as expected, the presence of counterweights causes the first natural frequency of the *4B-M* to considerably decrease (about 20%) with respect to the unbalanced mechanism *4B-U*. Moreover, two new mode shapes having a relatively low value of the natural frequencies appear: they become the second and the third mode, respectively. On the contrary, the natural frequencies of the other vibrational modes show a light increase with

respect to the *4B-U*. This last result, although not completely unexpected, differs from the common assumptions concerning the effects of counterweights on the modal parameters of a system. The modal behaviour of the *4B-U* and *4B-E* is highly similar: the differences among the modal parameters of these two models are negligible and no stiffening effect produced by the elastic elements may be appreciated. The reason is that the stiffness of the springs required to obtain gravity compensation is a few orders of magnitude lower than the stiffness of the other mechanism parts. Reasonably, this is expected to occur for most closed-loop mechanisms. Therefore, even in this case, an assumption often encountered in the literature may not be generally considered as well founded.

As for the KED simulations, Figs. 3–5 show the driving torque for the three mechanisms as a function of time, during the first four crank revolutions at the velocity of 1, 5 and 10 Hz, respectively. Table 2 reports the peak and RMS values of the torque of the mechanisms for all the working conditions (calculated for one full-speed cycle). Two remarkable results may be observed. First, gravity compensation is actually beneficial only for the lower crank velocity. In fact, as the angular velocity increases, the peak and the RMS values of the driving torque in the *4B-M* become significantly higher than in the *4B-U*, whereas those in the *4B-E* do not present relevant variations with respect to the unbalanced case. Indeed, at the low crank velocities, the inertia forces and the gravity loads are comparable: for the gravity compensated mechanisms *4B-M* and *4B-E*, a lower driving torque is therefore required. Second, for the highest crank velocity, the driving torque of all mechanisms is characterized by oscillations at the first natural frequency. In this case too, the *4B-U* and the *4B-E* behave similarly and the amplitude of the oscillations is rather small. Conversely, the oscillations of the driving torque in the *4B-M* are much more evident: in this case the vibrational effects appear to be rather severe, potentially penalizing the dynamic performances of the mechanism. Hence, the introduction of counterweights for static-balancing purposes may seriously worsen the functionality of the linkage. It is worth noticing that the mechanisms exhibit this resonant effect for a crank velocity that is about 20-25 times lower than the first natural frequency (depending on the specific linkage and its configuration). The study of the joint reaction forces showed other interesting results that will be discussed in a future work.

Table 1. Mean value of the first natural frequencies of the simulated four-bar linkages.

	Natural frequency [Hz]				
<i>4B-U</i>	262			1304	1983
<i>4B-M</i>	208	505	1091	1381	2034
<i>4B-E</i>	262			1304	1983

Table 2. Peak and RMS values of the driving torque.

Crank velocity	Torque [Nm]					
	1Hz		5 Hz		10 Hz	
	RMS	peak	RMS	peak	RMS	peak
<i>4B-U</i>	0.041	0.064	0.239	0.776	0.943	3.474
<i>4B-M</i>	0.013	0.047	0.336	1.211	1.481	5.961
<i>4B-E</i>	0.009	0.031	0.224	0.789	0.931	3.485

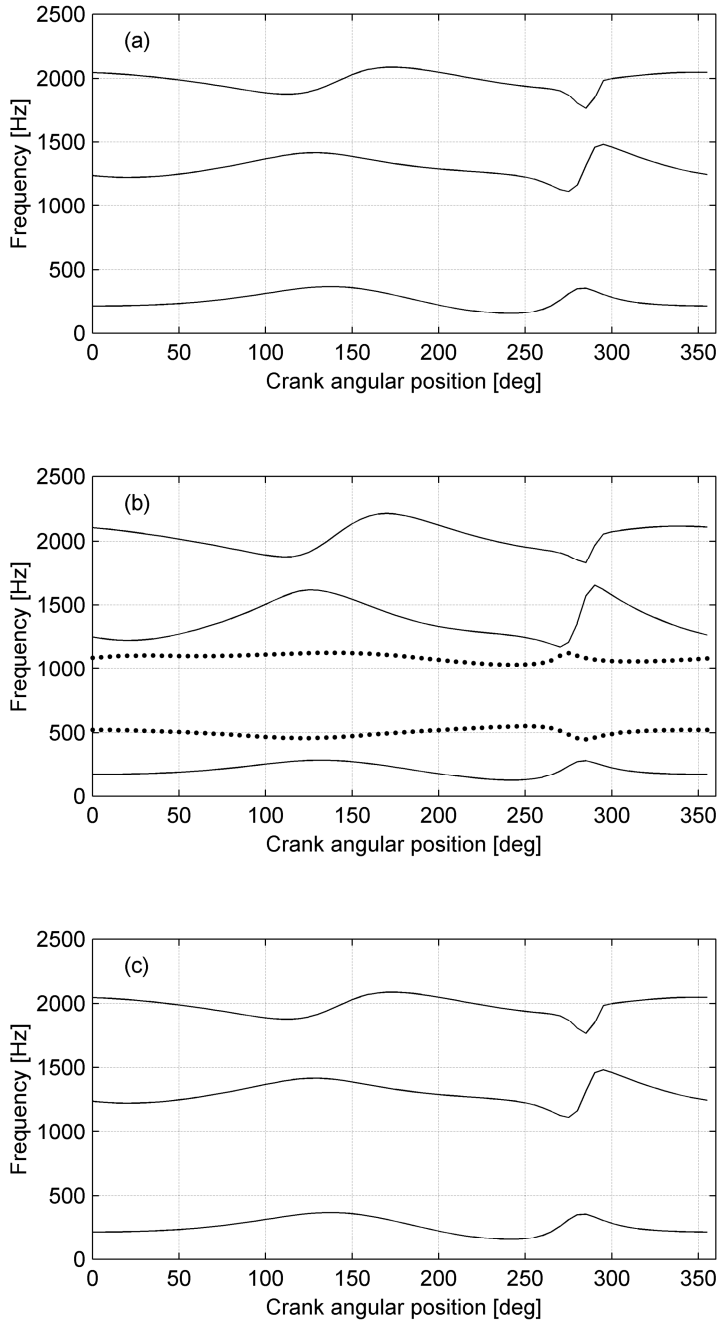


Figure 2. Natural frequencies of the four-bar linkages as a function of the crank angular position: (a) *4B-U*; (b) *4B-M*, (c) *4B-E*.

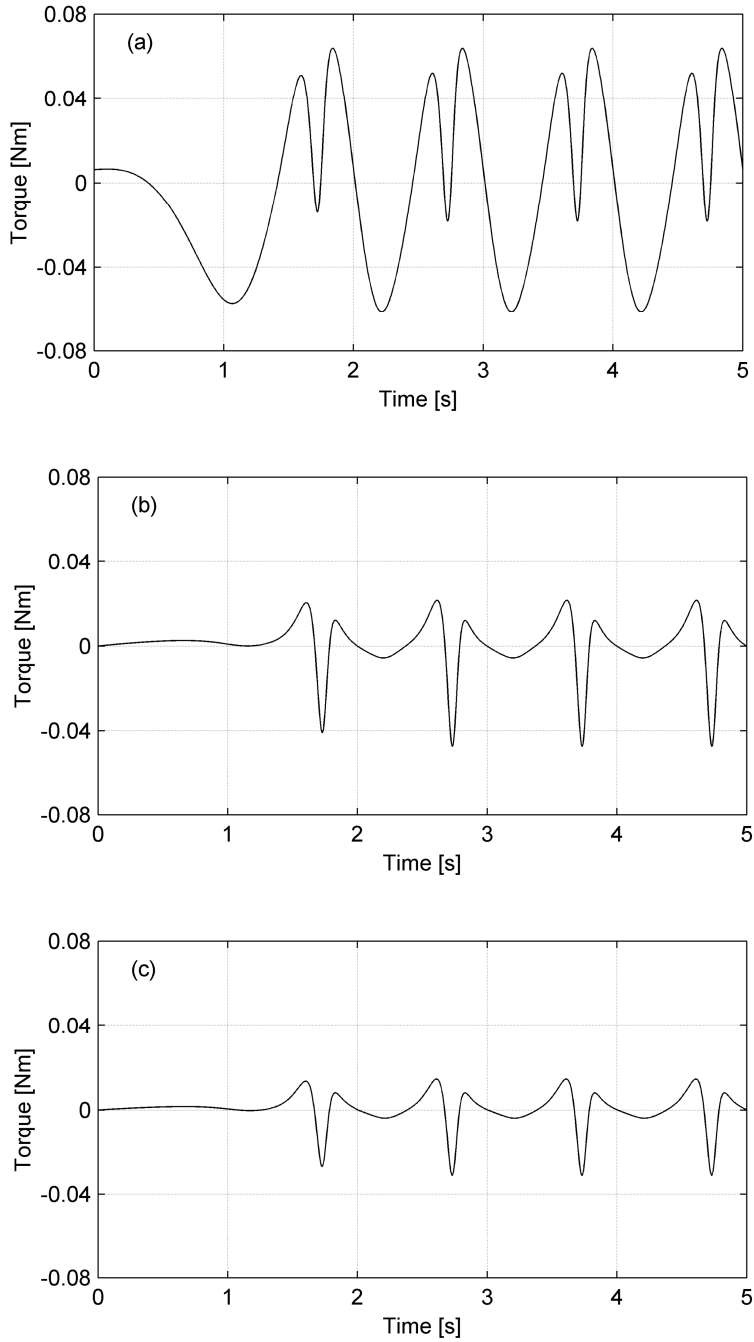


Figure 3. Driving torque of the four-bar linkages at the crank velocity of 1 Hz (for the first four crank revolutions): (a) *4B-U*; (b) *4B-M*; (c) *4B-E*.

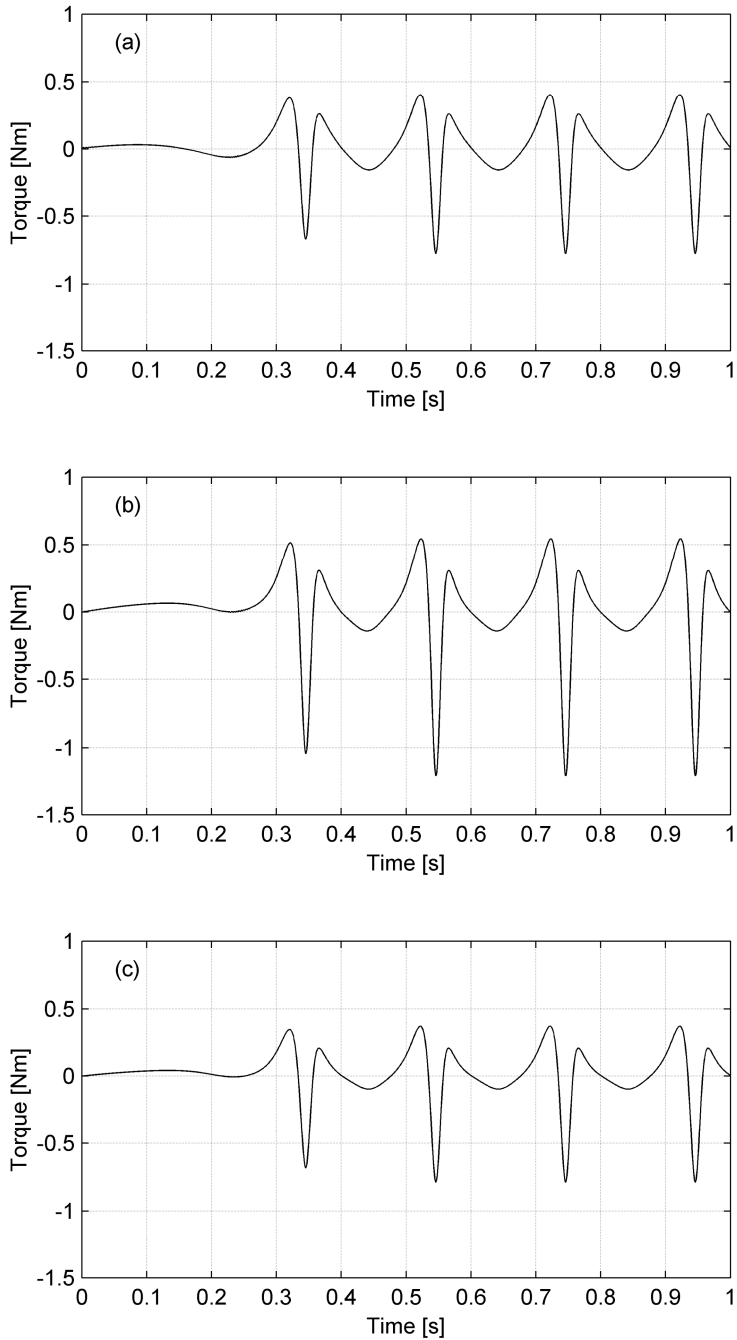


Figure 4. Driving torque of the four-bar linkages at the crank velocity of 5 Hz (for the first four crank revolutions): (a) *4B-U*; (b) *4B-M*; (c) *4B-E*.

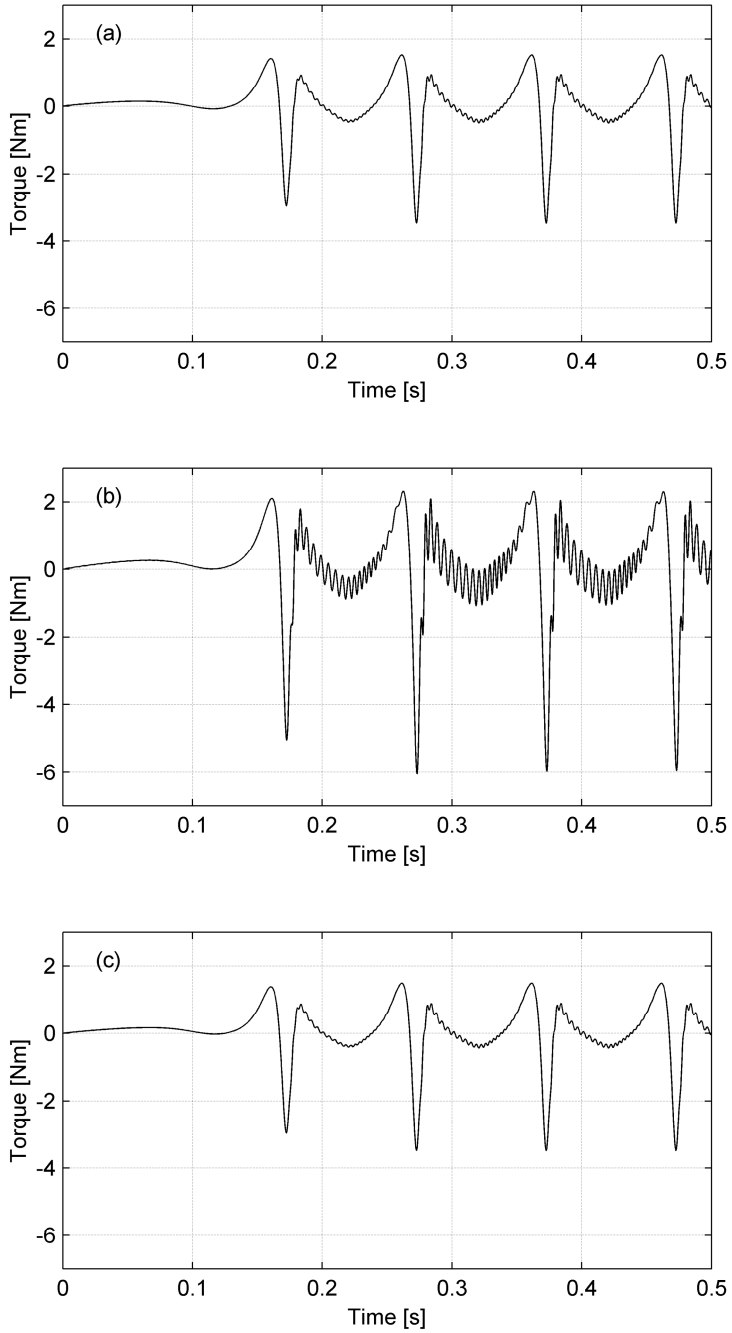


Figure 5. Driving torque of the four-bar linkages at the crank velocity of 10 Hz (for the first four crank revolutions): (a) *4B-U*; (b) *4B-M*; (c) *4B-E*.

4. CONCLUSIONS

The investigation of the effects of two different static-balancing techniques, namely mass balancing and elastic compensation, on the elastodynamic behaviour of a planar four-bar linkage was carried out; the compliance of all the links of the mechanism and that of the actuated joint were taken into account.

The results of the simulations showed that the introduction of counterweights heavily affects the dynamic behaviour of the linkage: as expected, the global stiffness of the mechanism decreases and a higher torque is required at medium/high velocities (i.e. as soon as the inertia actions become more relevant than gravity loads). Moreover, the mass-balanced mechanism shows an evident resonant effect for crank velocities significantly lower than the first natural frequency. Hence, the static balancing of a closed-loop mechanism by means of counterweights does not seem to be particularly convenient, due to the severe worsening of its dynamic performances.

The elastic compensation strategy does not affect the dynamic behaviour of the linkage. In particular, contrary to an assumption often encountered in the literature, the introduction of balancing springs does not produce any significant stiffening effect. It must be mentioned, however, that elastic compensation does not provide some advantages intrinsically associated with mass balancing, such as the balancing of the shaking force and the compensation of gravity loads for any orientation of the mechanism. Therefore a hybrid technique could represent an optimal balancing strategy.

According to these results, the presented topic of research seems to be worthy of further investigation. Moreover, the analysis may be extended to more complex closed-loop mechanisms, in particular spatial parallel manipulators, which, so far, do not seem to have been investigated from the elastodynamic point of view.

REFERENCES

- [1] Rivin, E.I., 1988. *Mechanical design of robots*. Section 3.5. McGraw-Hill, New York, 1988
- [2] Lowen, G.G., Tepper, F.R., Berkof, R.S., 1974. "The quantitative influence of complete force balancing on the forces and moments of certain families of four-bar linkages". *Mechanism and Machine Theory*, **9**(3-4), pp. 299-323.
- [3] Walker, M.J., Haines, R.S., 1982. "An experimental study of the effects of counterweights on a six-bar chain". *Mechanism and Machine Theory*, **17**(6), pp. 355-360.
- [4] Raghu, E., Balasubramonian, A., 1990. "Experimental Study on the Elastodynamic Behavior of the Unbalanced and the Counterweighted Four Bar Mechanisms". *Journal of Mechanical Design*, **112**(3), pp. 271-277.
- [5] Carricato, M., Gosselin, C., 2009. "A Statically Balanced Gough/Stewart-type Platform: Conception, Design and Simulation". To appear in the *ASME Journal of Mechanisms and Robotics*.
- [6] Lee, S.B., Cho, H.S., 1992. "Dynamic characteristics of balanced robotic manipulators with joint flexibility". *Robotica*, **10**(6), pp. 485-495.
- [7] Mahalingam, S., Sharan, A.M., 1986. "The optimal balancing of robotic manipulators". *IEEE International Conference on Robotics and Automation*, San Francisco, CA, pp. 828-835.

- [8] Xi, F., Sinatra, R., 1997. "Effect of dynamic balancing on four-bar linkage vibrations". *Mechanism and Machine Theory*, **32**(6), pp. 715-728.
- [9] Zobairi, M.A.K., Rao, S.S., Sahay, B., 1986. "Kineto-elastodynamic balancing of 4r-four bar mechanisms combining kinematic and dynamic stress considerations". *Mechanism and Machine Theory*, **21**(4), pp. 307-315.
- [10] Jandrasits, W.G., Lowen, G.G., 1979. "The elastic-dynamic behavior of a counterweighted rocker link with an overhanging endmass in a four-bar linkage". *Journal of Mechanical Design*, **101**(1), pp. 77-98.
- [11] Balasubramonian, A., Raghavacharyulu, E., 1987. "Influence of balancing weights on vibratory response of the coupler of a four bar mechanism". *Proceedings of the 7th IFToMM World Congress*, Seville, Spain, pp. 447-450.
- [12] Yu, Y.-Q., Jiang, B., 2007. "Analytical and experimental study on the dynamic balancing of flexible mechanisms". *Mechanism and Machine Theory*, **42**(5), pp. 626-635.
- [13] Berkof, R.S., Lowen, G.G., 1969. "A new method for completely force balancing simple linkages". *ASME Journal of Engineering for Industry*, **91B**(1), pp. 21-26.
- [14] Shieh, W.B., Chen, D.Z., Lin, P.Y., 2007. "Design of statically balanced planar four-bar linkages with base-attached springs". *Proceedings of the 12th IFToMM World Congress*, Besançon, France.

SIMULATION OF GEAR PUMP DYNAMIC BEHAVIOUR BY AN INTEGRATED LP/FE MODEL

Emiliano Mucchi

*EnDIF – Engineering Department in Ferrara,
Università degli Studi di Ferrara, Italy
E-mail: emiliano.mucchi@unife.it*

Giorgio Dalpiaz

*EnDIF – Engineering Department in Ferrara,
Università degli Studi di Ferrara, Italy
E-mail: giorgio.dalpiaz@unife.it*

Abstract. *This paper addresses external gear pumps for automotive applications, which operate at high speed and low pressure. It presents a hybrid lumped-parameter/finite-element model aimed at simulating the pump dynamic behaviour and evaluating the vibration produced on the case external surface. The developed model can be a very powerful tool in order to evaluate the vibration effect of different operational conditions as well as design improvements for noise and vibration reduction. As examples of application, in this paper the model is firstly used in order to evaluate the influence on the case and gear vibration of some parameters of the pumped fluid, as oil bulk modulus and viscosity, important for both dynamic behaviour and pump performance. Furthermore, the model has been modified in order to account of a design modification aimed at reducing pressure pulsations in the meshing area and therefore controlling noise and vibration.*

Keywords: *gear pump, pressure ripple, gear vibration, lumped-parameter model, FEM.*

1. INTRODUCTION

Gear pumps use a very simple mechanism to generate flow and consist of a low number of parts. They combine good performance and low costs. The simplicity of their design translates into higher reliability as compared to other positive displacement pumps using more complex design. However gear pumps are often accompanied with noise levels that are generally higher than other types of pumps. Such noise levels are the consequence of the dynamic forces within the system, related with the flow ripple inducing a pressure ripple as well as the variable meshing stiffness and gear errors. Thus, in the hydraulic systems using this pump as a power source, taking countermeasures for noise and vibration reduction is one of the key points.

In literature, several authors have addressed efforts to reduce noise and vibration in gear pumps by means of numerical and experimental analyses. References [1,2] contain a good description of the mathematical models used to simulate the meshing stiffness phenomena; the majority of the simulations to date are based on rigid-body lumped-parameter models with the meshing teeth assimilated to a lumped variable stiffness. On the other hand, the estimation of the pressure ripple as well as the pressure forces acting on gears is fully described in [3-6].

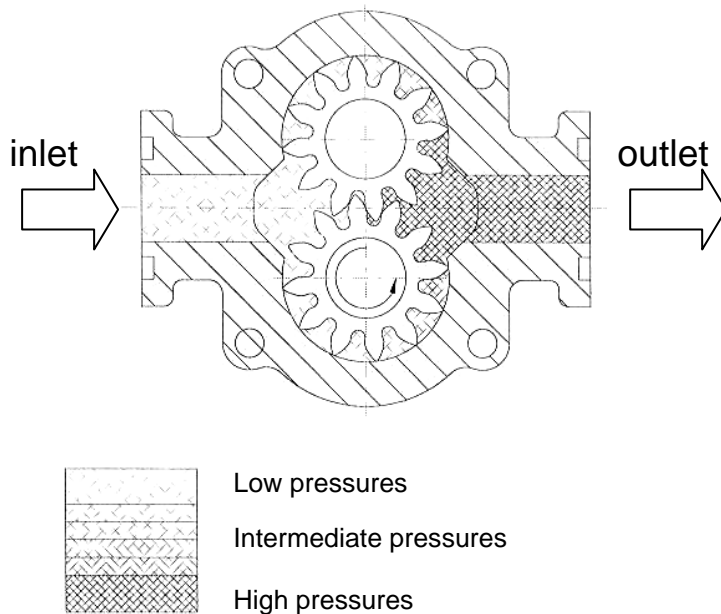


Figure 1. Cross-sectional view of an external gear pump.

The mechanical system under study is an external gear pump for vehicle steering. Figure 1 and Figure 2(a) show a cross-sectional view. In this kind of machine the principle of operation is very intuitive: the pump has two identical gears that are used for displacing fluid from the inlet to the outlet side around the peripheries. The gears are contained in a close-tolerance housing (external case) and are assembled by a couple of lateral floating bushes, hydraulically balanced, acting as seals for the lateral ends and as supports for gear shafts by means of two hydrodynamic bearings. The gears, the floating bushes and the case, enclosing all the components, define the isolated spaces that carry the fluid from the low to the high pressure chamber. Power is applied to the shaft of one of the gears - gear 1 in Fig. 2(a) - which transmits it to driven gear 2 through their meshing. The driving gear is connected to an electrical drive by an Oldham coupling.

In the gear meshing area, when two tooth pairs come in contact, a trapped volume could arise and could undergo a sudden volume reduction and consequently a violent change in its pressure. To avoid this, the trapped volume is put in communication with the high or low pressure chambers. That is the role of the relief grooves (Figure 2(a)) milled in the internal face of the lateral bushes whose shape and dimension are very important [7,9] in the resulting dynamic behaviour.

The authors have already developed and experimentally assessed a hybrid lumped-parameter/finite-element model of an external gear pump for automotive applications [7,14]; the lumped-parameter (LP) model regards the internal parts (rotors, gears, bushes), while the finite element (FE) model regards the external parts of the pump (case and end plates). The LP model is a non-linear kineto-elastodynamic model and includes the most important phenomena involved in the pump operation, as time-varying oil pressure

distribution on gears, time-varying meshing stiffness, gear errors and hydrodynamic journal bearing reactions [7,11,12]. The FE model makes it possible to carry out a forced vibration analysis for the evaluation of the acceleration levels produced on the external surfaces of the case by the excitation forces acting on the internal surface (bearing reactions and pressure forces), previously estimated using the LP model. In this sense the model is globally a hybrid LP/FE model. The model has been assessed using experiments: the experimental accelerations measured during run-up tests have been satisfactorily compared with the simulated accelerations coming from the LP/FE model [7,14]. As far as the authors are aware, models including the dynamic behaviour of gear pairs, the pressure evolution in a gear tooth space during the pump rotation, the oil bearing behaviour as well as the dynamic response of the external case were not found in the literature.

In previous works of the authors [7,9], the LP model was used in order to analyse the influence of a few design and operational parameters on the pump dynamic behaviour. In particular, the effect of operational pressure and speed, the influence of the clearance in the journal bearing and between tooth tip and pump case, and the effect of the dimension of the relief grooves in the bushes have been thoroughly discussed. As model results, this analysis gives the gear accelerations as a function of the above-mentioned design and operational parameters. However, the manufacturing companies and the customers are more interested in the evaluation of the accelerations on the external surface of gear pumps during operational conditions; in fact the vibrations on the external surfaces are directly related to noise and vibration transfer to the neighbouring structures (e.g. car chassis). As an example, in this paper, the hybrid LP/FE model has been used in order to evaluate the influence on the case vibration of a few parameters of the pumped fluid, as the oil viscosity and the bulk modulus, having a significant effect on the gear pump dynamic behaviour as well as pump performance. In fact, as the fluid viscosity increases, the amount of leakage flows (also called internal slippage) on the clearances among pump components decrease, making the pump more efficient, with flow rate approaching the theoretical value of the pump. Conversely, as the viscosity decreases, the amount of internal slippage increases and the flow rate reduces. On the other hand, the oil bulk modulus represents the resistance to uniform compression; for hydraulic fluids its value ($B_{Oil,p}$) depends on pressure and air fraction dissolved in the oil as follows [13]:

$$B_{Oil,p} = B_{Oil} \cdot \frac{1 + \frac{V_{Air}}{V_{Oil}}}{1 + \left(\frac{p_{ref}}{p} \right)^{\frac{1}{\gamma}} \cdot \frac{V_{Air}}{V_{Oil}} \cdot \frac{B_{Oil}}{\gamma \cdot p}} \quad (1)$$

where:

$B_{Oil,p}$: actual bulk modulus depending on pressure and air fraction in the oil mixture;

B_{Oil} : bulk modulus of pure oil at reference condition (atmospheric pressure);

V_{Air} : volume of the air at reference condition;

V_{Oil} : volume of the pure oil at reference condition;

p_{ref} : reference (atmospheric) pressure;

p : actual pressure;

γ : index depending on the process (it is equal to 1.4 in the case of isentropic process).

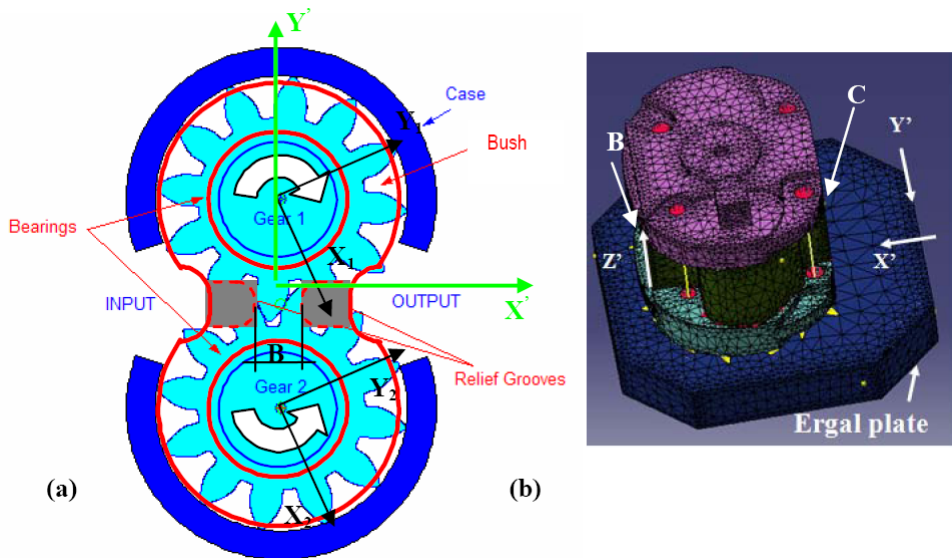


Figure 2. (a) Schematic of the gear pump and reference frames (Y-direction is parallel to the line of action and X-direction is perpendicular); (b) FE model with indication of reference frame $X'Y'Z'$ and response points B and C on the external case surface.

Moreover this hybrid model can be an useful tool for design optimization in the early phase of prototype development as well as in the following phase of design improvement and solution of functional problems.

The structure of the paper is the following. First authors give an overview of the hybrid LP/FE model highlighting the peculiar features as well as the available outputs of the model (Section 2). Afterwards, the attention shifts on the model results (Section 3) in terms of gear and case accelerations as a function of the fluid parameters (bulk modulus and oil viscosity). Then, Section 4 shows an example of application to design improvement in the early phase before the hard prototype is available: the model is modified in order to take account of a design modification aimed at reducing pressure pulsations in the meshing area and therefore controlling vibration and noise. In particular, the effect of additional grooves (called control grooves) milled in the internal face of the two lateral bushes has been included in the model. The acceleration results are given in the frequency domain as well as in the time domain (time synchronous average). Finally, Section 5 draws the conclusions.

2. OVERVIEW OF THE HYBRID LP/FE MODEL

The hybrid LP/FE model of the gear pump is aimed at evaluating the acceleration on the external parts of the gear pump (external surface of the case, flange, cover) during operational conditions. Hereafter, the LP and FE model will be described, but the complete model formulation and implementation can be found in the previous works of the authors [7,14].

The LP part of the hybrid model is a kineto-elastodynamic model aimed at evaluating the dynamic behaviour of the internal components of the gear pump. This LP model takes into account only transversal plane dynamics; this assumption is acceptable because model regards spur gears. The model has six degrees of freedom - for each gear, degrees of freedom are the displacements along directions X and Y in Figure 2(a) and the angular displacement - and includes the main important phenomena involved in the pump operation as time-varying oil pressure distribution on gears, time-varying meshing forces and hydrodynamic journal bearing reactions.

The meshing phenomena have been accurately modelled [7,11]; in fact, this type of gears works at low loads and consequently the effects of tooth profile errors and tooth separation could not be ignored. The model takes into account the parametric excitations due to the time-varying meshing stiffness and the tooth profile errors (obtained by a metrological analysis); in addition, the effects of the backlash between meshing teeth, the lubricant squeeze and the possibility of tooth contact on both lines of action were included in the model. The pressure distribution around the gears was obtained by means of a mathematical model with control volumes [7,15]. This model considers as control volumes the isolated tooth spaces “sealed” between teeth, lateral plates and case as well as the inlet and outlet chambers and the volume trapped between meshing teeth (called trapped volume). Applying to each control volume the continuity equation [7], the pressure variation, caused by the volumetric flow rate gain (difference between the flow rate coming into the volume and coming out) and by the volume variation during the gear rotation, can be computed by means of an iterative solution process. Then, the resultant pressure forces and torques are obtained taking into account the case wear due to the run in process, as well as the displacement of the lateral floating bushes. The non-linear behaviour of the hydrodynamic journal bearings is also included, as well as the torsional stiffness and damping of the driving shaft [7]. As a consequence, the model is highly non-linear.

The equations of motion [7,11] give the dynamic equilibrium among the pressure forces and torques on the gears, the meshing forces, the driving external torque, the bearing reactions and the gear inertia forces and torques. The contribution of the case vibration in the gear equilibrium is neglected in the LP model. In order to reduce the integration time, the average positions of the journal axes into the bearings are previously estimated. Thus, the variable pressure forces on gears can be approximately estimated in these average axis positions, before starting the numerical integration, obtaining an important reduction of the integration time.

The LP model was validated using experiments. In fact, a good model should be a good representation of the real machine and therefore should have a good correspondence with the experimental measurements. Tests were carried out on a test bench available at TRW Automotive Italia SpA – Divisione Automotive Pumps. The pump under testing is fastened on an ergal plate that also provides proper connections to low and high pressure oil pipes; moreover the plate is equipped with four high impedance quartz-based triaxial force sensors. The pump is driven by an electrical motor equipped with inverter having maximum angular speed of 5000 rpm; the plate has a hole for the driving shaft; the test bench operation is controlled by a system based on software Labview. The assessment of the LP model is shown in [7,10], where the comparison between simulations and experimental results concerning forces and moments has been presented; a specific validation procedure has been developed because it has not been possible to directly obtain vibration data concerning rotating components.

Table 1. Gear data concerning GENB pump.

Number of teeth	12
Gear module	1.15 mm
Face width	12.1 mm
Pressure angle	20°
Contact ratio	1.32
Base radius	6.484 mm

The LP model gives as output the dynamic forces and torques acting on the gears; in particular, the forces exchanged between the gears and the case are the pressure forces and bearing reactions. These are the forces exciting the case vibrations. Then, an FE model of the pump case (also including the experimental apparatus), Figure 2(b), was developed in LMS Virtual.Lab and MSC.Nastran and used in order to estimate the frequency response functions (FRFs) between the internal exciting forces and the vibrations of the external case surface. Finally, the actual acceleration of the case surface is evaluated in different operational conditions by multiplying the above-mentioned FRFs by the spectra of the exciting forces coming from the LP model.

The FE model not only regards the gear pump, but also the plate to which the pump is fastened during the experimental tests. In fact the model must be validated using experiments and therefore it has been modelled the same apparatus used for the experimental measurements, i.e. the pump, the ergal plate and the four force sensors supporting the plate as shown in Figure 2(b). Obviously, once the model has been assessed, it should be possible to model the real boundary conditions as in the actual location in a vehicle. It can be noted that the test bench used in the validation of the LP/FE model is the same above-mentioned test bench used for the validation of the LP model.

The FE model includes all the non-rotating components of the pump - i.e. the case, the two end-plates, the valves, the bushes - as well as the ergal plate to which the pump is fastened on the test bench and the four force sensors supporting the plate. The masses of the two gears are not included in the FE model because their inertial contribution is already taken into account by means of the dynamic forces computed by the LP model. The case, the two end-plates and the ergal plate are meshed by using tetrahedral elements as shown in Figure 2(b). In the actual pump, the case and the two end-plates are fastened together by means of two steel screws. The two screws are modelled as beam elements having the same cross-sectional area and inertia properties of the actual screws, joined to the surrounding mesh of the end-plate by means of interpolation spiders; this way, the motion of the two end-points of each beam element is defined as the weighted average of the motions of the surrounding mesh points. Furthermore, rigid spider connections are used in order to connect the case surface with the end-plate surfaces. The ergal plate is connected to ground by four triaxial force sensors located under the ergal plate. These sensors are modelled by means of spring elements. The model damping has been estimated using data coming from an experimental modal analysis (EMA).

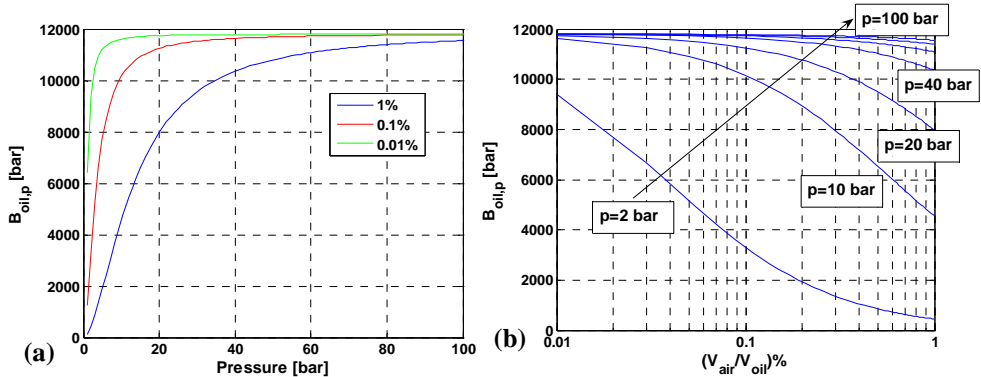


Figure 3. (a) Oil bulk modulus as a function of pressure for different air fractions (1%, 0.1%, 0.01%) and (b) oil bulk modulus as a function of air fraction for different pressure (from 2 to 100 bar).

The hybrid LP/FE model was then assessed using experiments: in particular, two linear run-up tests were performed from 2000 to 3350 rpm, with operational pressure of 34 bar and 90 bar using the above-mentioned test bench available at TRW, the same test bench used for the validation of the LP model. Case vibration was measured in points B and C of Figure 2(b), located on the external case surface at the output and input side, respectively. The waterfall maps of the acceleration spectra measured during the run-up tests were compared with the simulated ones coming from the hybrid model, relative to the same external points. In this phase, with the aim of better match the experimental results, some model parameters are adjusted; more details can be found in [7,10].

3. EFFECTS OF FLUID PROPERTIES ON PUMP VIBRATION

Hereafter the model has been used in order to evaluate the dynamic response of the system in terms of gear and case accelerations as a function of the oil bulk modulus and viscosity. The simulation results concern a pump with spur gears, namely GENB, whose main geometrical characteristics are summarised in Table 1. All the simulations have been performed at operational pressure of 90 bar and rotational speed of 3400 rpm for 40 complete rotations of the gears in order to have a long time history for averaging the results.

The influence of the bulk modulus on gear and case accelerations is firstly considered. The fluid used in the GENB pumps is Pentosin CHF11S, having bulk modulus B_{oil} of 11800 bar at atmospheric pressure and temperature of 80° C (reference conditions). On the base of Eqn. (1), the actual bulk modulus in operational conditions depends on pressure and air fraction in the oil mixture, as shown in Figure 3. In particular at low pressure (0-20 bar) the bulk modulus strongly depends on both pressure and air fraction, on the contrary at high pressure (60-100 bar) these bulk dependences are very slight and the bulk modulus approaches the value of pure oil at reference conditions (i.e. 11800 bar).

The output pressure of the pump under study can range between 3.5 to 100 bar; moreover, in the inlet side and in the trapped volume the pressure can reach very low values, close to the atmospheric pressure or sometimes lower (more details are given in [7,15]).

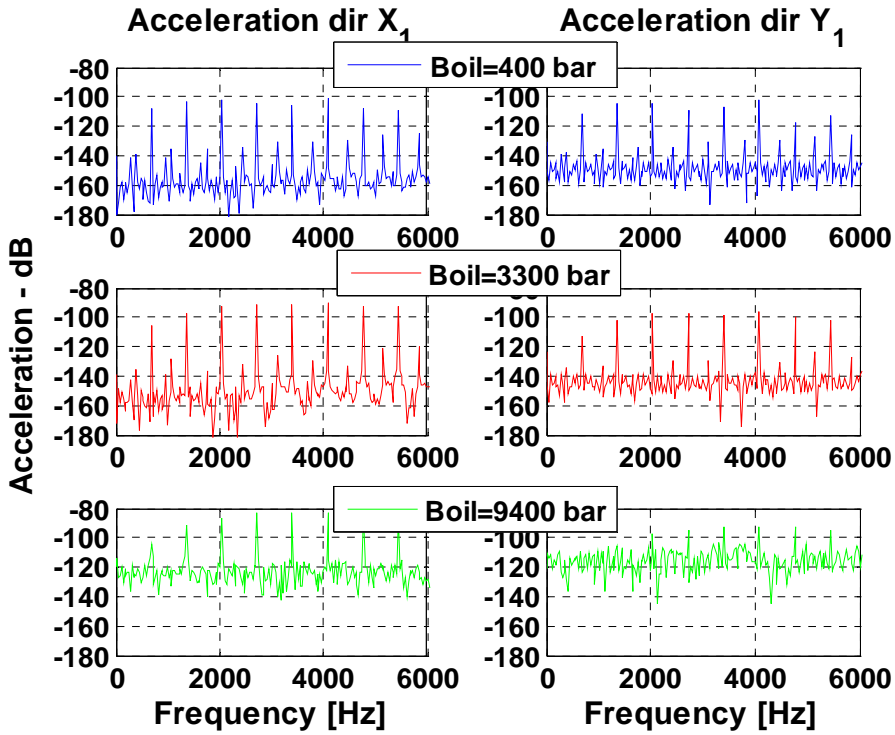


Figure 4. Spectrum of the acceleration TSA of gear 1 in direction X_1 and Y_1 , at operational conditions of 3400 rpm and 90 bar with different oil bulk modulus (400, 3300 and 9400 bar) and viscosity of 14 mPa.s.

Since a certain amount of air is always present in the operational oil, the actual bulk modulus of the oil contained in a tooth space varies during the rotation of the gears when the tooth space moves from the inlet to the outlet side. Depending of the actual air fraction and operational pressure, the bulk modulus can range between very low values (200-300 bar) till very high values (10000-11000 bar). In particular, the actual minimum value of the bulk modulus at the inlet side strongly depends on the air fraction in the oil mixture (Fig. 3(b)), while at the outlet side the influence of the air fraction on the bulk modulus is strong only for low operational pressure.

To date, the model cannot take account of the variation of the bulk modulus in the control volumes due to the pressure variation from the inlet to the outlet side: the bulk modulus is taken as constant in all the control volumes. In spite of this, the model results for different values of the bulk modulus are shown in the following, in order to give a general information about the sensitivity of gear and case vibrations to the bulk modulus. In particular, the considered values of bulk modulus are 400, 3300 and 9400 bar. On the base of these results, it will be evaluated if it is proper to update the model in order to take account of the bulk modulus variation in the control volumes.

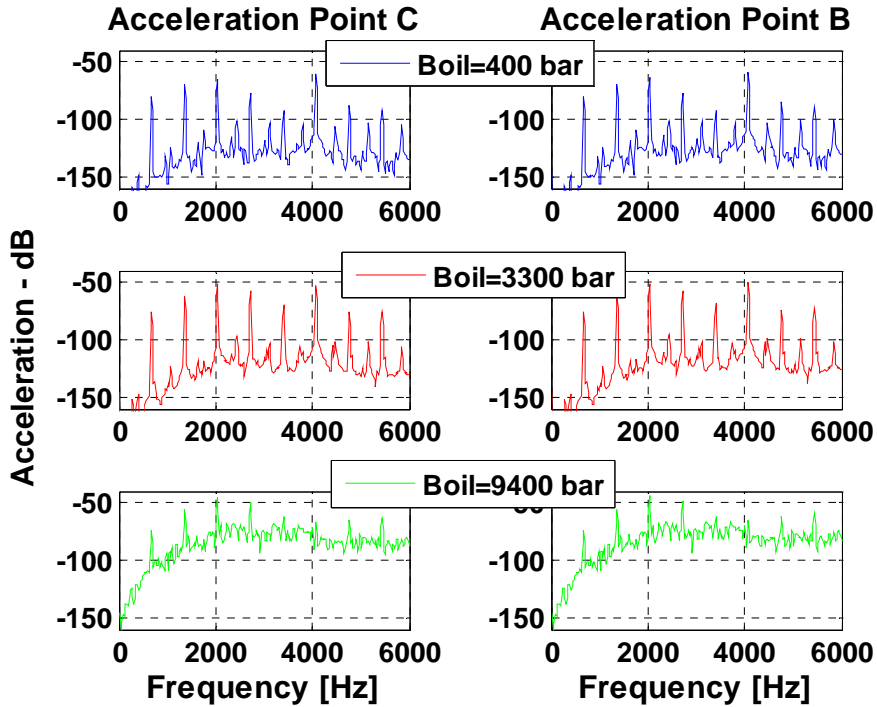


Figure 5. Spectrum of the TSA of the case acceleration in points C and B in direction X', at operational conditions of 3400 rpm and 90 bar with different oil bulk modulus (400, 3300 and 9400 bar) and viscosity of 14 mPa·s.

Figure 4 and Figure 5 illustrate the influence of the bulk modulus on the acceleration spectrum of gears and case, respectively. In particular, the case acceleration has been calculated in X' direction in points B and C of the external surface, shown in Figure 2(b). In order to smooth the results, the time synchronous average (TSA) over the available 40 rotations is firstly performed and the Fourier Transform of the TSA is then computed. Such acceleration spectra (as all the others in this paper) are displayed in the frequency range till 6 kHz and in dB value, with reference value of 10^{-6} m/s^2 .

In both gear and case spectra, the most important components are the meshing frequency and its harmonics, as expected. Moreover the higher the bulk modulus is, the higher the spectrum amplitude becomes. Therefore it seems that in order to reduce vibration, one has to choose oil with high air fraction. However, the designers have to face cavitation problems that can occur in this kind of pump [7,15]; in order to reduce the risk of cavitation in the inlet side, low air fractions are suggested.

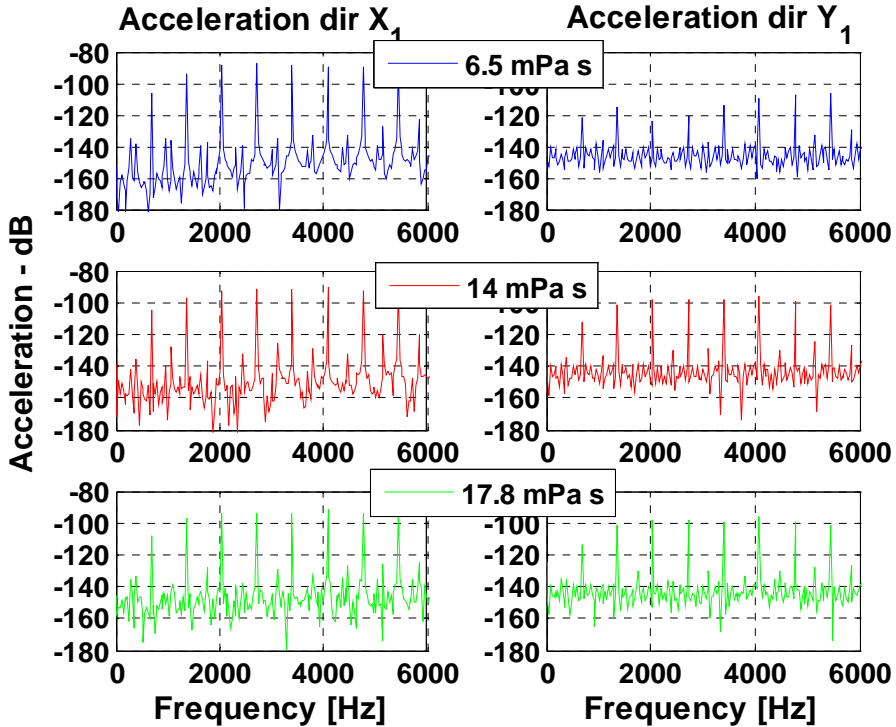


Figure 6. Spectrum of the acceleration TSA of gear 1 in direction X_1 and Y_1 , at operational conditions of 3400 rpm and 90 bar with different oil viscosity (6.5, 14 and 17.8 mPa·s) and bulk modulus of 3300 bar.

Comparing the corresponding gear and case acceleration spectra of Figs. 4 and 5, it can be noted that the meshing harmonics in the gear spectra are of similar amplitude, while the case spectra exhibit a progressive attenuation below 2 kHz and the components at about 2 kHz and 4 kHz are generally higher than the other ones. Moreover the case components are higher than gear ones of about 40 dB in the range above 2 kHz. This behaviour is obviously due to the case dynamic behaviour, acting as a “filter” for the gear accelerations that propagate from the internal to the external of the pump. Such a filter is actually the frequency response function of the case, that was estimated by the frequency analysis performed by means of the FE model of the case [7,14]: the case exhibits natural frequencies at about 2 kHz and 4 kHz [7], this explaining higher amplitude for the case components around such frequencies.

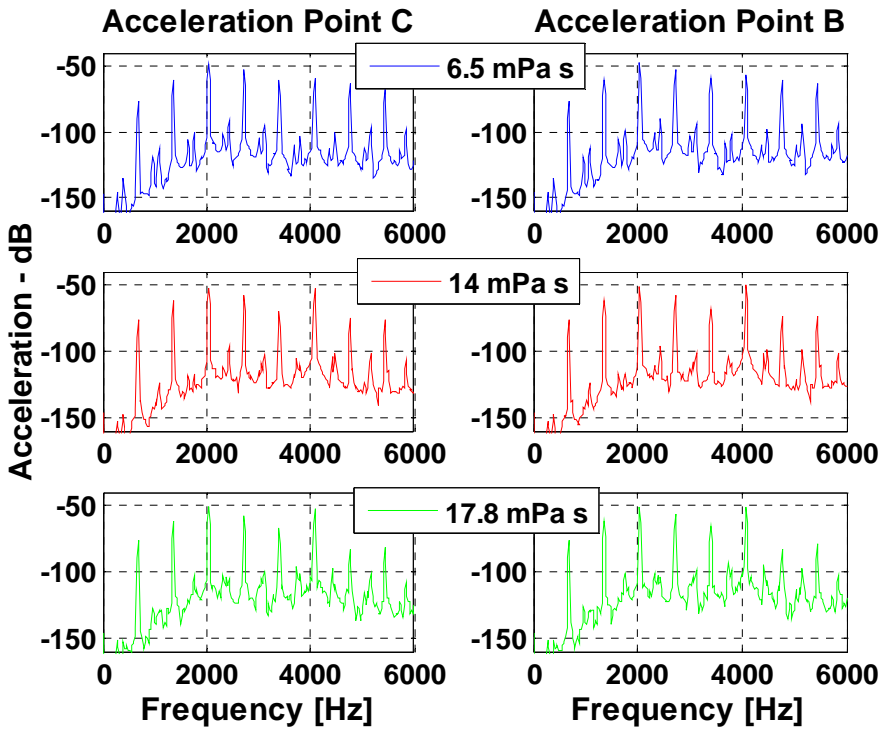


Figure 7. Spectrum of the TSA of the case acceleration in points C and B in direction X', at operational conditions of 3400 rpm and 90 bar with different oil viscosity (6.5, 14 and 17.8 mPa·s) and bulk modulus of 3300 bar.

Let now consider the influence of the oil viscosity. The oil viscosity strongly depends on temperature and more weakly on pressure in the range of the pump operational conditions. In standard operational condition, the oil of the GENB pump reaches the temperature of about 50°C, but it can range about 20°C to 80°C. It is interesting to evaluate the dynamic behaviour of the pump in terms of gear and case acceleration as a function of oil viscosity: the range 6.5 to 17.8 mPa·s has been chosen, considering the range of temperatures and output pressures in which the pump can work. Figure 6 and Figure 7 display the effect of variation of the oil viscosity in the gear and case acceleration, respectively. The effect of the oil viscosity on both gear and case acceleration is less important than the one of the bulk modulus, however differences occur in particular in direction Y_1 (Figure 6) where the peaks of the meshing harmonic are higher for higher values of oil viscosity.

Comparing the corresponding gear and case acceleration spectra of Figs. 6 and 7, analogous observation can be obtained, as comparing Figs. 4 and 5. As above, for case acceleration, the meshing harmonics at about 2 kHz and 4 kHz play a very important role in the overall vibration level of the pump case (see Figure 7).

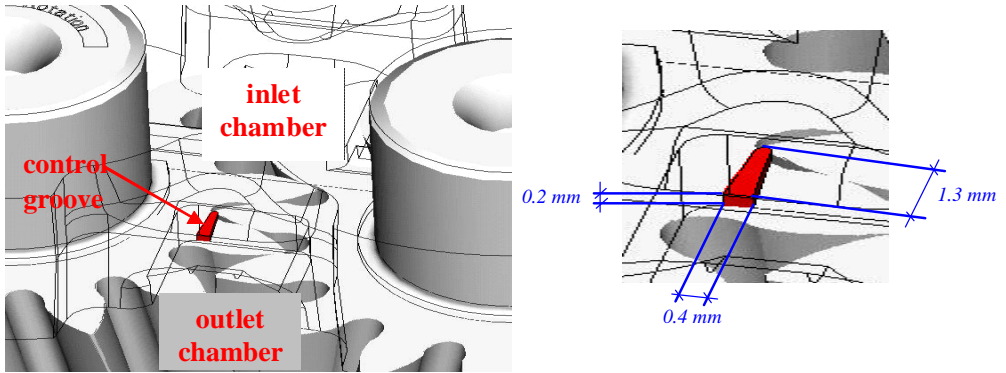


Figure 8. Control groove in the bush (left) and its dimensions (right).

4. EXAMPLE OF APPLICATION TO DESIGN IMPROVEMENT: EFFECT OF CONTROL GROOVE

To reduce the pressure pulsation in the trapped volume, producing high vibration levels, and taking under consideration efficiency aspects, the relief grooves have been redesigned, adding a further groove, called “control groove”, in each bush at the outlet side. Such a control groove (CG) aims at reducing the pressure increase in the trapped volume due to its volume reduction occurring during tooth meshing. In more details, the oil can leak in through the CG from the trapped volume to the output chamber, determining a pressure reduction and consequently low vibration level.

The control groove defines a new profile of the relief grooves as can be seen in Figure 8. The CG effect is an increased connection area between the trapped volume and the output chamber. This increment in the connection area is only present when the additional connection through the CG occurs, depending on the length of the CG itself. In fact, a minimum length allowing this connection exists and it depends on the tooth width, as shown in Figure 8. When there is connection through the control groove, the total connection area A_{out} becomes the sum of two terms:

$$A_{out} = A_{out,old} + A_{CG} \quad (2)$$

where $A_{out,old}$ is the connection area without the CG, which is variable depending on the position of the teeth in the meshing area with respects to the relief grooves, and A_{CG} is the additional connection area due to the CG.

Different control groove geometries have been considered. This Section compares the simulation results concerning the original bush design without the control groove and the design modification with a specific geometry of the control groove: length 1.3 mm, height 0.4 mm and thickness 0.2 mm (see Figure 8(b)). It is worth noting that the CG length of 1.3 mm taken in the simulation is longer than the above-mentioned minimum value allowing an effect in terms of flow rate and pressure pulsation.

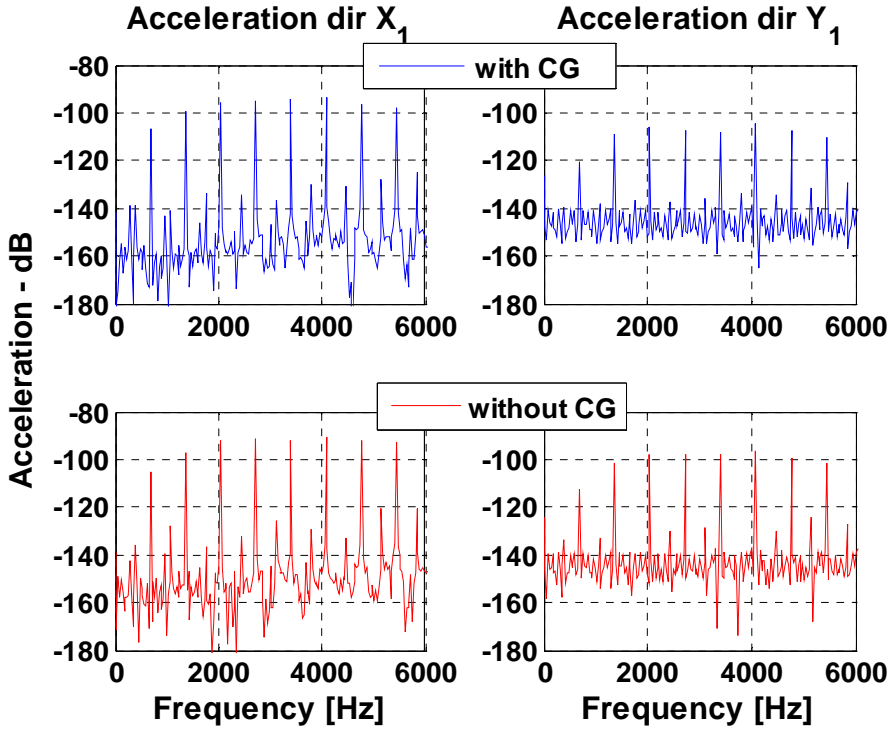


Figure 9. Spectrum of the acceleration TSA of gear 1 in direction X_1 and Y_1 , at operational conditions of 3400 rpm and 90 bar, with and without control grooves; bulk modulus of 3300 bar and viscosity of 14 mPa·s.

These CG dimensions lead to obtain the following increment of the connection area A_{CG} between the trapped volume and the output chamber: $A_{CG} = 0.4 \text{ mm} \cdot 0.2 \text{ mm} = 0.08 \text{ mm}^2$.

The pressure formulation of the LP model has been modified as follows. The CG effect has been introduced in the LP model as an increased connection area of the quantity A_{CG} . The flow rate between the trapped volume and the outlet chamber ($Q_{T,t,out}$) is calculated in the model assuming turbulent flow by using the following formulation [7,15]:

$$Q_{T,t,out} = \begin{cases} +0.65 \cdot A_{out} \cdot \sqrt{\frac{2 \cdot (p_t - p_{out})}{\rho}} & p_t > p_{out} \\ -0.65 \cdot A_{out} \cdot \sqrt{\frac{2 \cdot |p_t - p_{out}|}{\rho}} & p_t < p_{out} \end{cases} \quad (3)$$

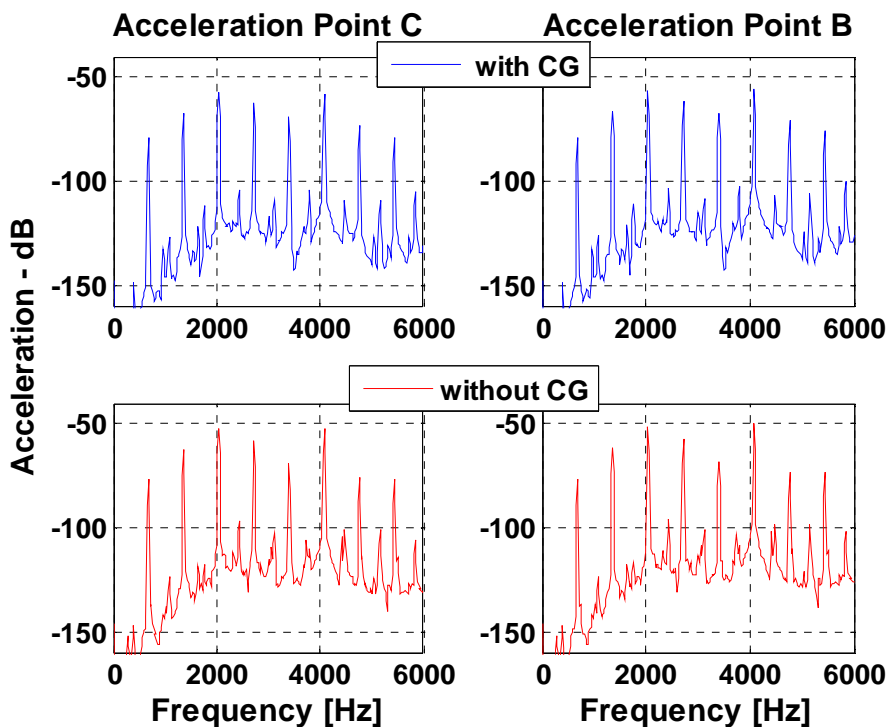


Figure 10. Spectrum of the TSA of the case acceleration in points C and B in direction X', at operational conditions of 3400 rpm and 90 bar, with and without control grooves; bulk modulus of 3300 bar and viscosity of 14 mPa·s.

where 0.65 is the flow coefficient depending on the Reynolds number, A_{out} is the total connection area between the trapped volume and the outlet chamber, ρ is the oil density and p_t, p_{out} are the pressures in the trapped volume and in the outlet chamber, respectively.

It is interesting to note that a variation in the connection area has a direct influence on the flow rate and an indirect influence on the pressure distribution around the gears as well. More specifically, a connection area modification (see Eqn. (3)) yields a change on the flow rate; such a flow rate change determines a modification on the pressure distribution (and therefore on p_t and p_{out}) that again modifies the flow rate value. Therefore, a local modification on the connection area determines a change on the entire pressure evolution around the gears.

Figure 9 and Figure 10 show the gear and case acceleration spectra with and without the contribution of the control grooves. The amplitude levels of both gear and case spectra are lower for the simulation with the CG effect: the amplitude of the meshing harmonics is reduced by adding the CG.

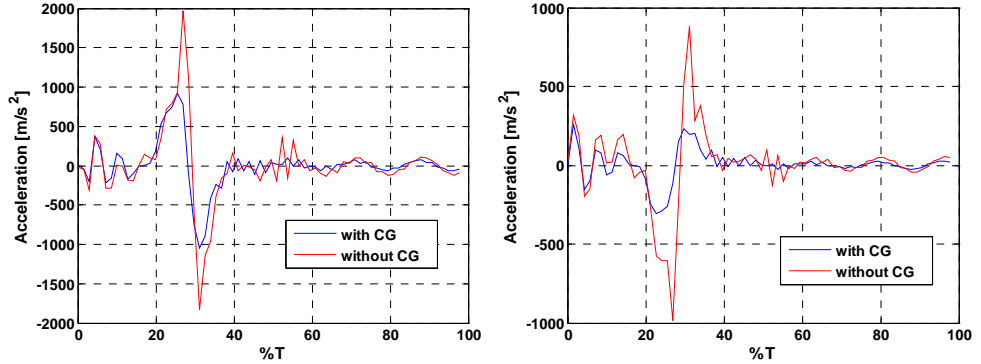


Figure 11. Acceleration TSA over 480 meshing periods of gear 1 in direction X_1 (left) and Y_1 (right), at operational conditions of 3400 rpm and 90 bar, with and without control grooves; bulk modulus of 3300 bar and viscosity of 14 mPa·s (T is the meshing period).

Moreover, it is worth noting that the meshing harmonics in the gear spectra of Fig. 9 are of similar amplitude in the range above 2 kHz, while in the case spectra of Fig. 10 the components at about 2 kHz and 4 kHz are generally higher than the other ones. This behaviour is due to the effect of the natural frequencies of the case, as already shown and discussed in Section 3.

Figure 11 shows the time synchronous average (TSA) of the gear accelerations over 480 meshing periods for Gear 1 in directions X_1 and Y_1 . The TSA is presented as a function of the meshing period T in percentage. The effect of the additional control grooves is very important: the amplitude of the acceleration peaks in the TSA are strongly reduced due to the introduction of the control groove. Thus the acceleration peaks can be reduced increasing the connection area A_{out} of 0.08 mm^2 by using control grooves of length 1.3 mm.

5. CONCLUSIONS

It is worth noting that the main operational problem in gear pumps is generally the high level of vibration and noise. In this context, a hybrid LP/FE model for the dynamic analysis of gear pumps for automotive applications has been developed and experimentally assessed. The model includes the dynamic behaviour of the gear pairs, the pressure evolution in a gear tooth space during the pump rotation, the hydrodynamic journal bearing reactions as well as the dynamic response of the external case. The lumped-parameter (LP) part of the hybrid model regards the internal components (rotors, gears, bushes), while the finite element (FE) part regards the external case and the end plates. The model is an applicative tool for manufacturing companies and customers interested in the evaluation of the vibrations on the external surface of gear pumps during operational conditions, which are directly related to noise and vibration transfer to the neighbouring structures (e.g. car chassis). The model can be used for evaluating the influence of operational parameters and fluid characteristics on the case vibrations. Moreover this hybrid model can be an useful tool for design optimization in the early phase of prototype development as well as in the following phase of design improvement and solution of functional problems.

In particular, this paper has shown two applications of the model: the evaluation of the

influence on the gear and case vibration of oil bulk modulus and viscosity, and the effect of a new design of the relief grooves in the bushes, with the introduction of additional control grooves. These simulations have brought to draw the following conclusions:

- ✓ In the range of the considered values, the higher the bulk modulus is, the higher the acceleration amplitude becomes for both the case and the gears.
- ✓ The effect of the viscosity in terms of gear and case acceleration is less important than the one of the bulk modulus, however differences occur in particular in the gear acceleration in direction Y_1 , where the acceleration peaks are higher for higher values of oil viscosity.
- ✓ On the base of these results it would seem useful to update the model, taking account of the variation of bulk modulus with the pressure evolution in the control volumes during gear rotation. This could increase the accuracy of the simulation results.
- ✓ As a design improvement, the introduction of a control groove of proper geometry in the bushes has been considered: it produces an important reduction in the amplitude of both gear and case acceleration. This is obtained by means of a control groove of length 1.3 mm, giving an additional area of connection between the trapped volume and the outlet chamber of 0.08 mm^2 .
- ✓ This hybrid model is effective in order to evaluate the contribution of the pump case that acts as a flexible body amplifying several harmonics of the meshing frequency.

Acknowledgements

This work has been developed within the laboratory of research and technology transfer InterMech (Division Acoustics and Vibration - LAV) realized through the contribution of Regione Emilia Romagna - Assessorato Attivita' Produttive, Sviluppo Economico, Piano telematico, PRRITT misura 3.4 azione A.

REFERENCES

- [1] Ozguven, H.N., Houser, D.R., 1988. "Mathematical Models Used in Gear Dynamics – A review". *Journal of Sound and Vibration*, **121**(3), pp. 383-411.
- [2] Velez, M., Maatar, P., 1996. "A mathematical model for analyzing the influence of shape deviations and mounting errors on gear dynamics behaviour". *Journal of Sound and Vibration*, **191**(5), April, pp. 629-660.
- [3] Mancò, S., Nervegna, N., 1987. "Modello matematico di pompe oleodinamiche a ingranaggi esterni", *Oleodinamica-Pneumatica*, **1**.
- [4] Mancò, S., Nervegna, N., 1993. "Pressure transient in an external gear hydraulic pump", Proceedings of the Second JHPS International Symposium on Fluid Power, T. Maeda, eds., 14 October 1993 Tokyo.
- [5] Paltrinieri, F., Dilani, M., Borghi, M., 2002. "Modelling and simulating hydraulically balance external gear pumps". Proceedings of the 2nd International FPNI Ph.D. Symposium on Fluid Power, 3-6 July 2002, Modena, Italy.
- [6] Borghi, M., Milani, M., Paltrinieri, F., Zardin, B., 2005. "Pressure transients in external gear pumps and motors meshing volumes". SAE International 2005-01-3619.

- [7] Mucchi, E., 2007. "Dynamic analysis of external gear pumps by means of non linear models and experimental techniques". PhD. Thesis, EnDIF-Engineering Department in Ferrara, Università degli Studi di Ferrara, Ferrara, Italy, March.
- [8] Canotti, L., 1998. *Fluidi Oleodinamici 2*, Cemoter-CNR, 496-1998.
- [9] Mucchi, E., Rivola, A., 2007. "Influence of design and operational parameters on the dynamic behaviour of gear pumps". Proceedings of the ICSV14, 9-12 July 2007 Cairns, Australia.
- [10] Dalpiaz, G., Fernández del Rincón, A., Mucchi, E. and Rivola, A., 2005. "Experimental validation of a model for the dynamic analysis of gear pumps", Proceedings of Novem 2005, April 2005, Saint Raphael, France.
- [11] Dalpiaz, G., Fernández del Rincón, A. and Mucchi, E., 2004. "Modeling meshing phenomena in gear pumps", Proceedings of ISMA 2004, Sas P. and De Munck M. , eds., September 2004, Leuven, Belgium.
- [12] Childs, D., Moes, H., Van Leeuwen, H., 1977. "Journal bearing impedance descriptions for rotordynamic application", *Journal of lubrication technology*, **99** pp. 198-214.
- [13] Murrenhoff, H., 2001. "Grundlagen der Fluidtechnik Teil 1: Hydraulik". Shaker Verlag, ISBN 3-8265-9446-0, pp. 121.
- [14] Mucchi, E., Venturi, V., Dalpiaz, G., 2007. "A hybrid LP/FE model for the dynamic analysis of external gear pump". Proceedings of the International Mechanical Engineering Congress & Expositions IMECE2007, 11-16 November, Seattle, Washington, USA.
- [15] Dalpiaz, G., Mucchi, E., D'Elia, G., Fernandez Del Rincon, A., 2006. "Pressure phenomena in dynamic analysis of external gear pumps", Sas P. and De Munck M. , eds., Proceedings of ISMA2006, 18-20 September, Leuven, Belgium.

FULL-FIELD ESPI TECHNIQUES IN STRUCTURAL DYNAMICS

Alessandro Zanmarini

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: alessandro.zanmarini@mail.ing.unibo.it

Abstract. *In this paper Electronic Speckle Pattern Interferometry (ESPI) technique is presented for the structural dynamics of mechanical components. Starting from a brief description of its principles, ESPI is shown with its potentials for the inquiry in complex-shaped displacement fields, due to heterogeneous structure or complex vibration patterns, in a wide frequency range. The applications range from integrity inspection, to vibration analysis and to fatigue life assessment. This paper illustrates a review of the researches of the author on the ESPI capabilities in experimental dynamics.*

Keywords: *full-field measurements, ESPI, displacement field, structural dynamics, vibrations, NDT, dynamic strains & stresses, fatigue*

NOMENCLATURE

bold array notation

λ wave length of coherent laser light

\mathbf{s}_i sensitivity vector in ESPI instrument

$\Delta\phi_i(x, y)$ phase map measured along \mathbf{s}_i

$\mathbf{d}(x, y)$ global displacement map

$\boldsymbol{\varepsilon}(x, y)$ strain tensor map

$\boldsymbol{\sigma}(x, y)$ stress tensor map

K_r, b fatigue strength coefficient and exponent

\mathbf{S}_r map of stress cycle ranges

\mathbf{S}_{eq} map of equivalent range of broad band stress cycles

1. INTRODUCTION

In the last decades the awareness of noise and vibration harshness (NVH) problems, the significant advances in structural refinement and the continuous demand of high mechanical performances have asked for more advanced analysis and measurement technologies, especially in automotive and aerospace engineering. Virtual modelling has strongly spread out in many fields limiting the test number, but asking for even more reliable benchmarks for its models. Tests have thus nowadays to cope with engineering complex problems and their related design technologies and tools. Optical techniques offer peculiar characteristics that make them suitable for tasks where established technologies, e.g. accelerometer or extensometer arrays

that are used with great success for a wide variety of other tests, become unsatisfactory regarding the expected performances of the designed component. Optical measurements are non-contact, which avoids the problems related to the added inertia of the transducers on the measured structure; in some tests, the relative motion of the parts may require the total absence of sensors. Another peculiarity of optical measurements is that a huge amount of experimental degrees of freedom (dof) may be explored synchronously, if optically accessible. The range of these measurements is extended down to microns and nanometres. On top of that, some techniques deliver 3D measurements and contouring data, allowing strain calculations with greater precision than by means of extensometric arrangements. When dealing with high dof number, optical techniques become competitive also in terms of cost per dof.

Digital Holography/ESPI is a full field technique and its direct competitor is another well established optical technology, Laser Doppler Vibrometry (LDV), which measures the velocity in a single point of a surface, rather than the entire surface displacement field. LDV works in the time/frequency domain, while ESPI in the image-space/frequency domain. ESPI accesses a great amount of dofs simultaneously, with no need of time-consuming scans, so giving rise to high definition and consistent displacement fields in a very short time; no surface preparation is generally needed. Together with contouring measurements and 3D displacement fields, ESPI helps to enhance the quality of the measurements on complex shaped objects, where errors about location and orientation of transducers are usually of great concern. The high quality and detail of the ESPI measurements could find application where increasing spatial refined experimental models are requested, like in FE model updating or hybrid modelling, flutter or squealing researches, modal strains recovery or where the high modal density gives rise to complex spatial patterns. Light weight structures NVH, panel acoustic radiation, high definition Operating Deflection Shapes (ODS) measurements, strains&stress calculation, noise control and modelling, fatigue life prediction, system identification, dynamic instability cancellation and many others, also diagnostics and Non Destructive Testing (NDT), can take advantage of ESPI peculiarities.

This paper is a review of the author's researches carried on this field. The researches were started in the framework of the European project *Speckle Interferometry for Industrial Needs* at Dantec Ettemeyer, Ulm, Germany, in 2004. They moved from the principles of ESPI in static fields, to assess proper vibration measurement procedure in complex-valued fields to reach FRF estimation and portability of the results toward other NVH environments. The application of a robust differentiation approach led to the evaluation of complex-valued dynamic strains and stresses, for the estimation of the fatigue life of components in broad band forced dynamics. The presented methodologies have been implemented in C-language softwares; the graphic interface, from pictures to animations, has been realized by means of OpenGL graphic libraries in the GNU-GCC environment.

2. ESPI PRINCIPLES

Speckles patterns in Fig.(1) are the result of interference phenomena made by monochromatic light when reflected by a surface with a micro-roughness comparable to the wavelength λ of the incident light [1,2]. ESPI technology employs lasers with high *coherence length* to insure the constant phase delay of the photons along the same light path in the testing set-up. The reflected wave is superimposed on a digital sensor with a reference wave and their sum is recorded in each exposure; the difference between two exposures gives the interferogram (or specklegram), with its typical fringe pattern. Fringes are isolines of phase difference between

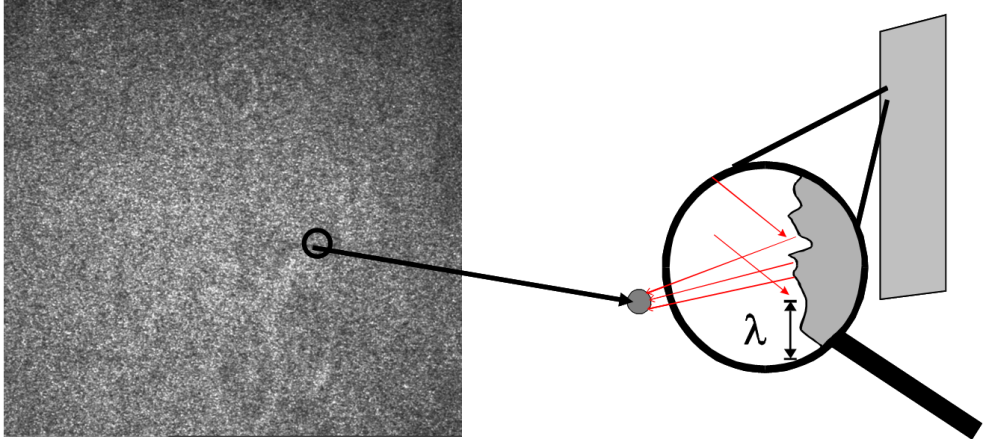


Figure 1: Origin of speckles

the records and correspond to optical path length differences of the beams. If no other cause is present, $\Delta\phi_i(x, y)$ is the deterministic phase difference map induced by surface movement between the two measuring steps as in Fig.(2). Different procedures can be used to quantitatively extract $\Delta\phi_i(x, y)$ from the specklegrams as in Fig.(3), based on phase shifting [3,4] or spatial FFT filtering [5] techniques, and later to design the sensors accordingly. Higher resolution in the digital sensor is reflected as an increased quality in solving the specklegrams, since it enhances the smooth transition of the grey levels from black to white inside each fringe, for the advantage of the phase unwrapping algorithms.

In Fig.(4) is sketched the optical scheme of the sensor and illumination unit. Defining the sensitivity vector s_i as the difference between the unit vectors of the illumination and the observation direction, the phase map $\Delta\phi_i(x, y)$ is linked to the projection of the object global displacement map $d(x, y)$ along the s_i as follows:

$$\Delta\phi_i(x, y) = \left(\frac{2\pi}{\lambda} \right) d(x, y) \cdot s_i \quad (1)$$

When s_i approximates the observation direction, a displacement of roughly one λ in the out of plane direction is covered by every fringe step of 2π in the interferogram. By using at least

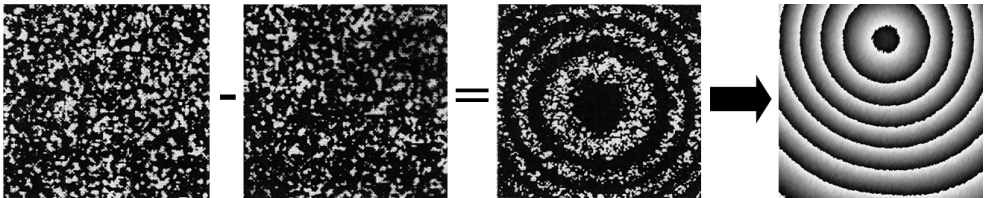


Figure 2: From exposures to specklegram/fringes and phase map

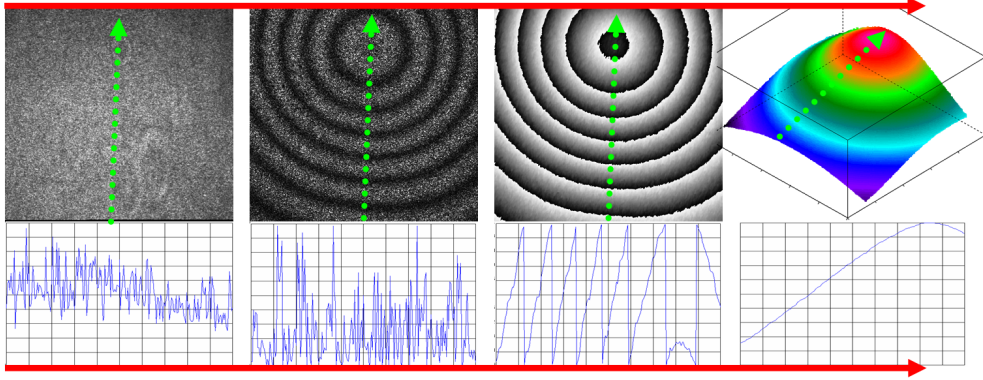


Figure 3: Displacement evaluation steps

three sensitivity vectors (one sensor - multiple light sources, or one light source - multiple sensors), 3D displacement maps between two exposures can be acquired as follows:

$$\begin{pmatrix} d_x(x, y) \\ d_y(x, y) \\ d_z(x, y) \end{pmatrix} = \frac{\lambda}{2\pi} \begin{pmatrix} s_{1x} & s_{1y} & s_{1z} \\ s_{2x} & s_{2y} & s_{2z} \\ s_{3x} & s_{3y} & s_{3z} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \Delta\phi_1(x, y) \\ \Delta\phi_2(x, y) \\ \Delta\phi_3(x, y) \end{pmatrix}$$

or

$$\mathbf{d}(x, y) = \frac{\lambda}{2\pi} [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3]^{-T} \cdot \Delta\phi(x, y) \quad (2)$$

In [3–5] is shown that ESPI can also be used to get the 3D shape of the surface, by means of piezoelectric controlled mirrors that vary the \mathbf{s}_i vectors while the sample is fixed. Shape measurement become relevant in strain calculations when the sample surface is manifold and not flat as in the works here recalled.

Examples of measuring systems

When measuring static deflections, the ESPI procedure consists in simply acquiring the exposures before and after the load application, with the best exposure timing to enhance the quality of the fringes. Dealing with dynamics, there is the extra need of an illuminating source that let the equipment capture still fringes on a moving object, being the latter undergoing a transient event or forced harmonic vibrations.

Pulsed ESPI. The measurement system here shown is the Dantec Ettemeyer Q600 Puls-ESPI, as in Fig.(5). Spatial FFT filtering techniques [5] are embedded in the system to evaluate the phase maps from the interferograms. The system is equipped with an Innolas Ruby laser (class 4), with $\lambda=694$ nano meters (red light), 1 meter coherence length. The total energy of the two pulses is of 1 Joule, released in few nano seconds. Red filters can be used in front of the sensor lenses to stop the blue and green components of day light; neutral density filters can also be used to attenuate the overall intensity in each camera, so as to balance the exposition level in each sensor and to enhance the interferogram quality. The delay between

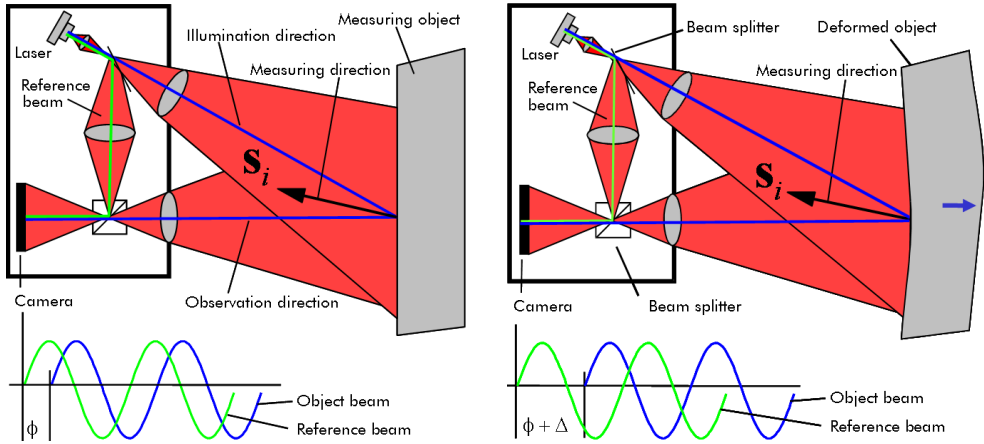


Figure 4: Optical scheme: reference state (a) and deformed state (b) exposures

the pulses, driven by computer control, can be selected in the range 2-800 μ s; alternatively, manual release is also possible in order to acquire the two needed interferograms. A triggering electronic device completes the equipment to deal with transient and harmonic events (frequency range of 20 Hz - 40 kHz).

The size of the measurable area is up to some squared meters when using appropriate divergent lenses (like a C-mount 15-75mm zoom lens). The three sensors use a 1300 x 1030 pixels CCD and are coupled to the laser through optical fibres. The system is able to measure displacements in the range 60 nm - 10 μ m; it follows that it provides extreme high sensibility to surface displacement. The distance of the system from the target, as well as the three sensor positions on their arms, must respect the coherence length condition (the difference between the reference beam and object beam path must be lower than the coherence length of the laser), and must be selected to obtain in-plane displacement sensibility, avoiding the singularity of the sensibility matrix in Eq.(2).

Stroboscopic ESPI. Dantec Ettemeyer Q500 VibroESPI with contouring long arms was used in these 3D single harmonic measurements. The system embeds phase shifting and contouring techniques to evaluate the phase maps from the interferograms. In Fig.(6) it can be seen the set-up for forced harmonic excitation and related Q500 measuring system. The use of long arms makes the sensibility matrix of Eq.(2) better conditioned, since the sensibility vectors form a bigger angle between one another, which means that the accuracy in the in-plane measurements is improved. Stroboscopic illumination was here generated. The single wavelength light was provided by a Coherent Verdi V5 cw Nd:YVO₄ 5 Watt laser (class 4), with $\lambda=532$ nm (green light), with more than 100 m coherence length. A continuous wave laser can be made stroboscopic by means of an acoustic/optical modulator, driven by the system internal triggering electronics at the same frequency as the excitation, and a spatial filter to select the first harmonic component as in Fig.(6). The stroboscopic laser beam was coupled to the sensor by means of adjustable mirrors fixed on the optical table. Inside the Q500, the coherent light source is driven by piezoelectric shutters and mirrors, to

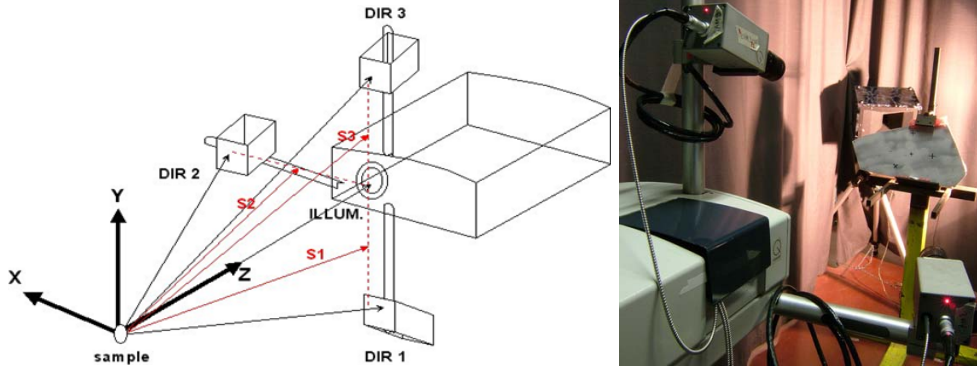


Figure 5: ESPI 3D displacement measurements with pulsed laser: a) scheme, b) system

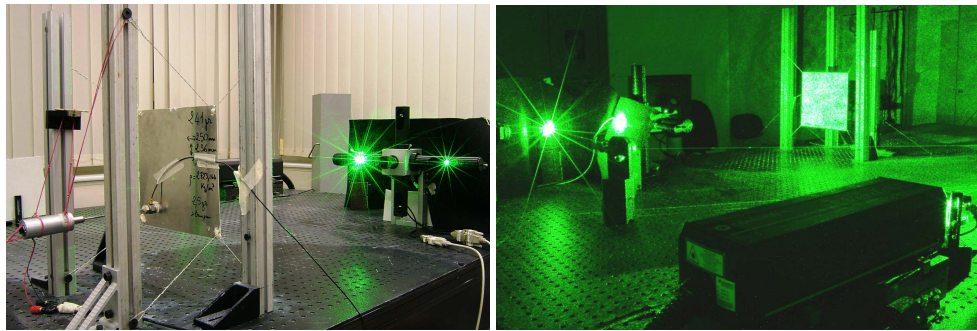


Figure 6: ESPI 3D displacement measurements with stroboscopic laser: laboratory set-up

cope with the embedded phase shifting technology [3, 4]. The CCD sensor had an array of 768 x 576 pixels. A 12-62 mm C-mount zoom lens equipped the cameras; the size of the measurable area in the frequency range 20 Hz - 40 kHz came up to a 0.75 m diameter, by using appropriate divergent lenses. The range of the measured displacement for this system is 30 nm - 3 μ m.

3. APPLICATION: NON DESTRUCTIVE TESTING [6]

Impact damage detection on composites is considered of uttermost importance for in-service delaminations and also matrix cracking and fibre fracture. A quality assessment of crucial components is therefore indispensable, since it is well known that detachments or delaminations between the layers (or fibres and matrix) of composites frequently occur with a high speed propagation and high degree of danger that generally can bring the part to its sudden collapse. It should be remembered that, especially for structural components, the monitoring actions have to be taken on site, in operational conditions and must be strictly non-invasive.

Vibration analysis based techniques and non-destructive testing (NDT) methods are well established in quality control and life predictions of crucial industrial components, but ex-

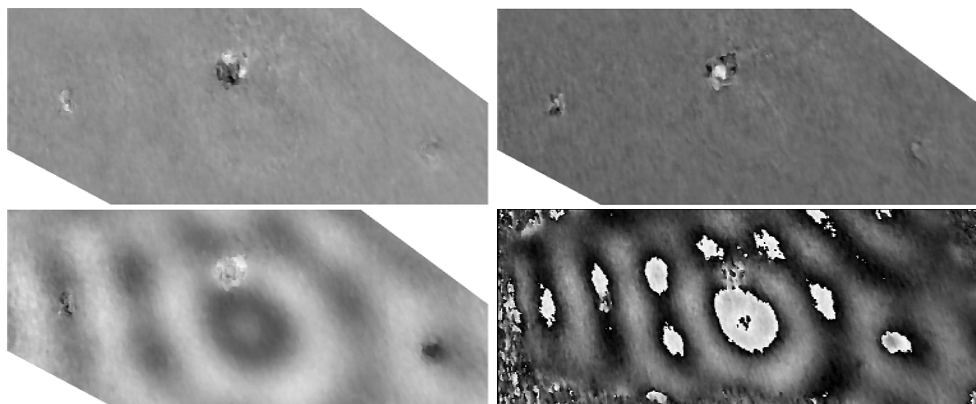


Figure 7: NDT results [6], acoustic excitation at 5600 Hz: a) dir-X, b) dir-Y, c) dir-Z, d) one sensor's phase map

perimental and operational modal analysis based approaches proved to be not sufficiently sensible to the structure alterations due to small damages.

ESPI takes advantage of the interference fringe systems as an encoded rich information source that can directly suggest hidden deterioration or surface flaws in a full field and highly detailed view of the component. A fringe singular and very local alteration shows an underneath inhomogeneity of the material structure from the neighbouring areas, that may be index of hidden defect. Detection and analysis of the different local behaviours in the measured fringe systems and displacement maps might become the key-issue in damage assessment, in judging, sizing and locating such imperfections inside the material. ESPI offers the advantage of full field optical non-contacting measurements by having access simultaneously to a great amount of experimental dofs without requiring any time-consuming scanning of a probe over all the areas, as happens in conventional ultrasonic inspection, eddy currents or SLDV; no structural alteration of the component under test is induced by the sensors. As it is material and load independent, ESPI measurement approach is quite a universal technique, while limitations might derive from the object displacement ranges. ESPI can be used for in-situ measurements with the aid of a pulsed high energy illumination source able to freeze nearly any ODS of the component surface in the interferograms. This leads to the measurement of the transient and single frequency excitation structural responses, and frees the inspections of the need of a vibration insulated table on which to locate the samples under test. Pulsed Speckle Interferometry was here tested as a measuring methodology able to supply the extremely high spatial resolution and displacement singularity sensibility expected for a successful damage assessment in composite structures in operational conditions. ESPI in this work can be considered environment independent, whereas modal approaches might be sensible to temperature and humidity changes, since neither reference nor model was required as the material singularities were searched by a qualitative analysis of displacement gradients on the structure surface and not by the comparison with an integer specimen. No extra data post-processing was here used, so that the procedure was suited for fast detections, nearly in real time.

Beside the measuring technique, a relevant role was played by the appropriate excitation

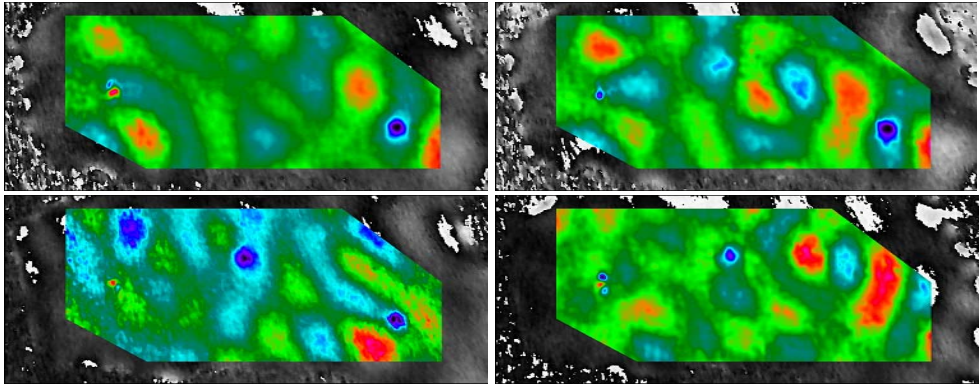


Figure 8: NDT results [6], Z-displacement patterns due to shaker excitation: a) 4750, b) 4800, c) 5600, d) 6200 Hz

approach able to give inhomogeneities a singular behaviour toward neighbouring material.

NDT prepared sample

The object under test was a 530 mm wide, 315 mm high and 40 mm thick sandwich panel. The cellulose fibre honeycomb core was bonded by araldite adhesive to two external skins in epoxy/fibreglass unidirectional cross-ply, laminated in four layers. Three known faults were inserted in the upper skin by wax disks, with a diameter of 30 mm, to simulate delaminations, voids and debondings. *Fault 1*, the most superficial one, was put between the first and second layer on the left of the fibreglass fabric; *fault 2*, being the deepest, was placed on a central location between the honeycomb core and the last fibre layer; *fault 3*, in an intermediate deep position, was between layer three and four on the right of the panel.

NDT excitation approaches

Four different loading approaches were investigated, as loading sources that give high displacement gradients. The first one consisted in exciting the panel with an acoustic emitter, as shown in Fig.(5b): in this way a high energy pressure field, modulated by a sine wave function generator and amplifier, was driven towards the back side of the panel thus exciting it contact-less at each selected frequency. The second approach was obtained through the dynamic load of an electrodynamic shaker, directly positioned on the back side of the panel, driven by sine wave excitation. The third load source was made by using a 1 kW halogen lamp on the front defected side of the panel; short time exposition was applied, so that thermal gradients were imposed only on the first layers of the sandwich skin, to have an effective singular displacement in the damaged areas. The last loading condition employed was the application of a few Newtons mechanical static load, on the right side boundary and with the normal direction of the panel, to superimpose a global bending over the vibration pattern induced by the shaker, which was acting at the same time. Also impact testing was performed to study the transient by hammering the panel, but the first global bending modes of the structure masked any local singular behaviour of the skin.

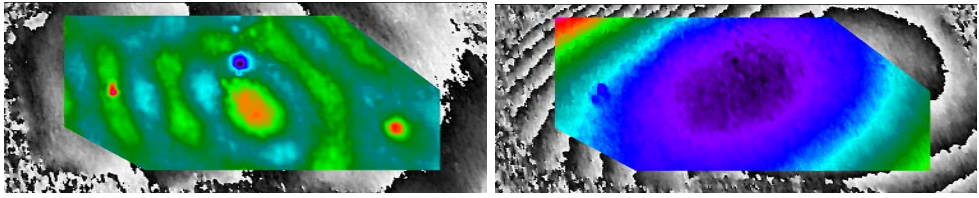


Figure 9: NDT results [6], a) combined shaker and acoustic excitation at 5600 Hz, b) thermal, 1 kW halogen lamp

NDT results

Depending on its nature and depth of location, each defect is generally sensible to the loading condition in a different way. Generally speaking, a basic rule might be suggested, i.e. that at low frequency the superficial defects are more evident, while at higher frequencies the deeper defects seem to be better revealed; and also that small defects become more detectable at high frequency, while big flaws are evident starting from static loads. At the same time it is important to keep in mind that damage assessment is a function of the overall structural dynamics of the component. This means that a defect may be detectable on some loading conditions, while being invisible on some others due to the eigenmode superposition.

With acoustic excitation (Fig.(7)) the flaw detection was clear enough from the out-of-plane displacement, coded either roughly in each one-sensor phase map or exactly in the calculated Z-displacement field (Eqn.(2)); the in-plane displacement maps confirmed even clearer the damage assessed by the out-of-plane information. It might be remarked that with acoustic excitation not all the three defects were depicted with the same outstanding clearness. At the same time this kind of excitation provided a frequency dependent and contactless loading source.

In the case of shaker excitation with much higher energy input, the assessment by means of the three-dimensional displacements gives clearly the location of all the flaws, as can be seen in Fig.(8a,b,d). In-plane displacement produced mainly uniform noise except in the defects. The defects might be otherwise mismatched with a feasible peak or the scattering of the fringe border in a single phase map, as in Fig.(8c). Adding a static load to the shaker action gave a clearer assessment of all the three defects as in Fig.(9a), since it was sufficient to evaluate either the out-of-plane or one-sensor information only.

Thermal excitation gave the panel a global static displacement shape able to reveal only *fault 1* (Fig.(9b)). There was a global change of the shape in all the directions and not only in the defects. The defect was clearer from the X-direction displacement map.

While *fault 1* is detected mostly in every test, from static to high frequency ones, giving rise to the conclusion that it might be quite superficial (as it was described above), the other two showed a frequency/energy dependent behaviour. It is thus not possible from these tests only to state a general rule regarding the type, depth and real size of the defects, such as a statistical study with known references and benchmarks might add.

NDT notes

A relevant advantage in damage assessment came from three dimensional full-field high sensitive measurements, since the singular in-plane displacements of the surface were precious

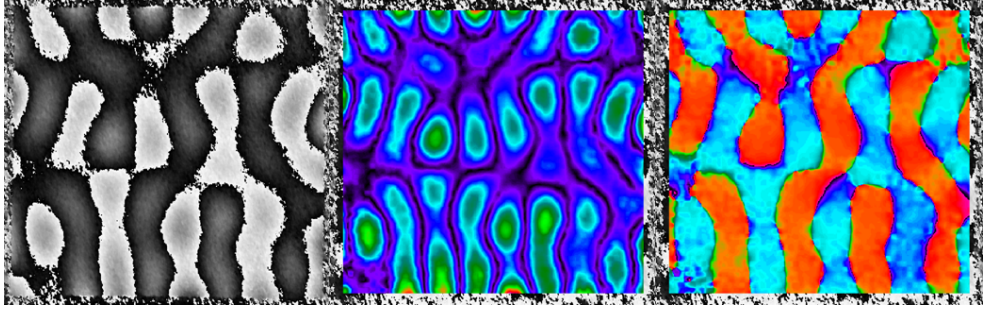


Figure 10: Dynamic ESPI evaluation (i -th phase map, complex amplitude & phase)

indexes of defect presence and location. Dynamic loading approaches gave better opportunities to reveal the structure singularities by means of inquiring their frequency and energy dependence, to be exploited in a systematic analysis regarding the relation between the print on the displacement maps and the real properties of the defect.

4. DYNAMIC ESPI MEASUREMENTS

The high stability of stroboscopic laser illumination makes this latter preferable for steady events, while pulsed lasers are generally used for unsteady events (but being more scattering in performance). Using single harmonic excitation and stroboscopic illumination, it follows that the moving surface is seen like a static deflection shape, function of the peculiar frequency, so a fringe system can be regularly evaluated.

Being complex-valued the nature of each ODS in the frequency domain, an ODS caught between two exposures will be a projection on real-valued domain, affected by the (adjustable) phase delay between the illumination and the excitation. There is thus a deep linkage between the displacement field extracted and the phase references chosen, and the results can greatly differ also on the same single frequency forced response, leading to shape misunderstandings. It is of great importance in dynamics have an approach that be independent from the phase delays taken from the excitation. A solution to this problem comes from searching for the complex-valued representation of the displacement fields as single frequency forced responses. The first step of the procedure consists in running n successive equally phase-shifted measurements by adjusting the acquisition timing. Each measurement is accompanied by sampled absolute references in one location of the measured field: this keeps consistent each couple of exposures with the next ones, which are not acquired on the same wave, due to instrumentation intrinsic speed, but on those delayed in time by an integer multiple of the wave period, finally regarding each displacement field as being sampled on the same cosine wave. The results are n displacement fields each of them shifted from the other by the known quantity $2\pi/n$; the i -th displacement map $\mathbf{d}_\omega^i(x, y)$ can be simply modelled by a threshold value $\mathbf{a}_\omega(x, y)$ and a cosine wave of amplitude $\mathbf{b}_\omega(x, y)$ with phase delay $\mathbf{c}_\omega(x, y)$, relative to the excitation:

$$\mathbf{d}_\omega^i(x, y) = \mathbf{a}_\omega(x, y) + \mathbf{b}_\omega(x, y) \cdot \cos(\mathbf{c}_\omega(x, y) + (2\pi/n) \cdot i) \quad (3)$$

Non-linear least squares fitting approach can then be used to extract the parameters ($\mathbf{a}_\omega(x, y)$,

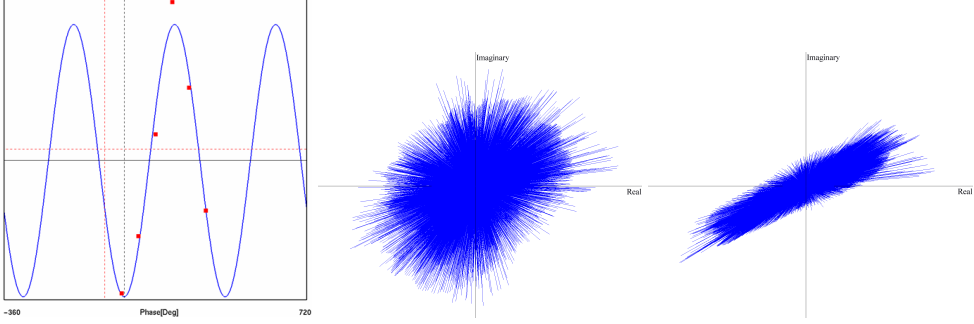


Figure 11: Cosine fitting of phase shifted responses (a) and complex-valued displacement dof distribution on the complex plane of two different ODSs (b,c)

$\mathbf{b}_\omega(x, y)$, $\mathbf{c}_\omega(x, y)$) of this n -data sequence in Eqn.(3). Disregarding the constant term, the complex-valued data (real and imaginary part) at the frequency ω of excitation become:

$$\mathbf{d}_\omega(x, y) = \{\mathbf{b}_\omega(x, y) \cdot \cos(\mathbf{c}_\omega(x, y)), \mathbf{b}_\omega(x, y) \cdot \sin(\mathbf{c}_\omega(x, y))\} \quad (4)$$

The evaluation of complex-valued data, as explained, delivers measurements that are independent of the phase parameters of the single acquisition; at the same time it releases to the experimenter the whole complex-valued nature of the displacement fields, whose advantages will be explained below. The fitting procedure has also a filtering effect on the noise in the harmonic signals. The described procedure is a way to allow ESPI to measure complex-valued data fields at a specific harmonic frequency excitation. Besides, it is also possible to obtain the complex-valued data of the input, by means of the same procedure on the simultaneously acquired external channels as in [7]; it becomes feasible to calculate a single line of a receptance FRF by complex-valued division of the output by the input at that specific frequency ω . In Fig.(11) are displayed six equally phase delayed steps (red squares), belonging to the force transducer channel, and fitted by the waveform (blue), together with two examples of the distribution on the complex plane of complex-valued displacement at a single frequency. A complete spectrum can then be obtained by repeating the procedure for a series of stepped sine sweeping frequency lines, like in Fig.(12), and packing together the obtained results, like at the dawn of modal analysis [8].

The importance of the complex-valued nature in data fields

In structural dynamics it is of uttermost importance have experimental data as complex-valued in the frequency domain. When dealing with models, it is well known that for general damped structures only a complex-valued representation of modes is consistent with the solution sought; any reduction to real-valued modes might introduce errors in figuring the physical problem in proportion to the misunderstanding of dissipative phenomena.

Many effective procedures are based on the complex-valued data obtained by transforming time-domain signals that generally lack of high spatial definition. ESPI delivers displacements in the frequency-domain only and not in the time- one; but having complex-valued frequency dependent data gives the chance to add the high spatial definition of ESPI to the effectiveness-proved investigation tools typical of structural dynamics. In the case where

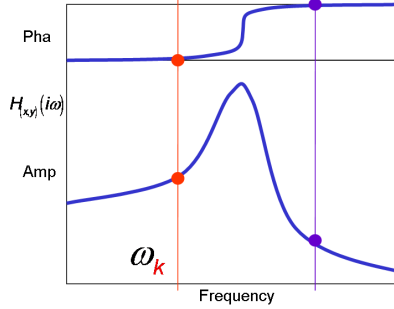


Figure 12: Frequency Response Function

many modes concur to the superposition in an ODS, complex-valued measurements reflect the phase difference between the modes, as function of the modal damping and eigenfrequency of each one. This is more evident in the transition from a frequency region where one single mode is predominant to other regions where a different shape is displayed: the superposition of the modes, which might give responses with comparable amplitude and variable phase delays toward the excitation, give the ESPI-ODS more the nature of a travelling wave pattern instead of a standing wave, with a high complexity in terms of imaginary part in the data that would be completely lost if a projection on real-valued domain be done, like projecting the data of Fig.(11b, c) on any single line.

Complex-valued ESPI might be of high value for dealing with general damping modelling in real life structures, with boundary conditions and anisotropies. Instead, real-valued fields can not reveal the inner nature of non-uniform phase delays, precluding relevant analyses on the behaviour of structured materials and products, leading to potential misunderstandings of physical problems.

5. APPLICATION: DYNAMIC ANALYSIS OF A BRAKE DISK [9]

Brake squeal noise occurs when the friction between the rotor and pads creates coupling of modes of various structures as dynamic instability. Past experimental approaches confirmed that the squealing frequency is near the individual eigenfrequencies of the components, and that the frequency of squeal does not depend on the rotational speed of the disc. Modes characterized by shapes with a high relative displacement at the contact point have the highest propensity to squeal.

The panorama outlined so far asks for accurate experimental tools able to cope with the dynamic characterization of brake system components in a wide frequency range and with high spatial definition. This latter plays a relevant role in detecting potential matching patterns in the dynamic behavior of single components, from the most global bending modes at low frequency to the very local cupping ones at high frequency. Rough sampling has to be avoided in order not to lose important information in assessing the tendency of modes to generate unstable vibrations when modal cross coupling occurs. A successful experimental session is the base for tuning the models from which to get a better understanding of the phenomenon and design the squeal cancellation, like mode decoupling, shifting or mode coupling strength reduction.

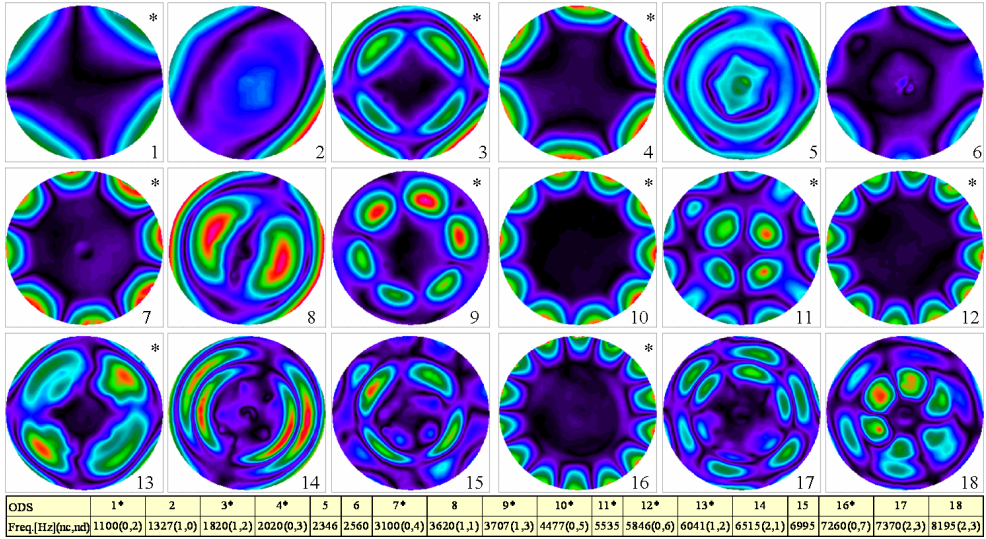


Figure 13: Brake disk dynamics [9] in the frequency range [1100-8195 Hz]

Dynamic measurement set-ups

The object under test was a commercial self-ventilating brake disc from the automotive market. The knuckle was fixed to a rigid frame in both set-ups. The excitation was provided in the out-of-plane direction by an electrodynamic shaker (with 30 kHz maximum frequency), and driven by a frequency function generator. Dantec Ettemeyer Q500 VibroESPI was used in these out-of-plane complex-valued displacement measurements.

Dynamic analysis considerations

ESPI techniques have the characteristic of providing displacement results in the spatial domain at selected frequencies rather than in a fully populated discrete frequency domain, as structural dynamics experimental tools, like accelerometers and LDV after Fourier transforms, usually do. ODS are different from eigenmodes since they sum up all the behaviours of the structure at that frequency; nevertheless, when acquired at resonances of single modes, single frequency ODS might be very close to code the proper eigenmode information. At the same time, the accurate study of the ODS behaviour when sweeping over a resonance frequency can be of value to analyse the doublet modes, their frequency shift or the nodal spatial distribution that derives from imperfect axisymmetry of the disc. Besides increasing the data spectral density in a wide range demands nowadays a long time for ESPI measurements. Nevertheless the latter can be driven in a semi-automatic way by selecting the number of spectral lines wanted in a frequency band; care must be taken to adjust the amplitude of the excitation in each band manually, so that to avoid noise in the phase maps due to a high number of fringes, which lead to irresolvable displacements. On the contrary, a high number of dofs data (up to the CCD pixel number) is contained in each single frequency ODS obtained from one ESPI measurement in few seconds. It follows that the dynamic behaviour of the complex shape is shown at each frequency line with an extreme spatial accuracy, providing

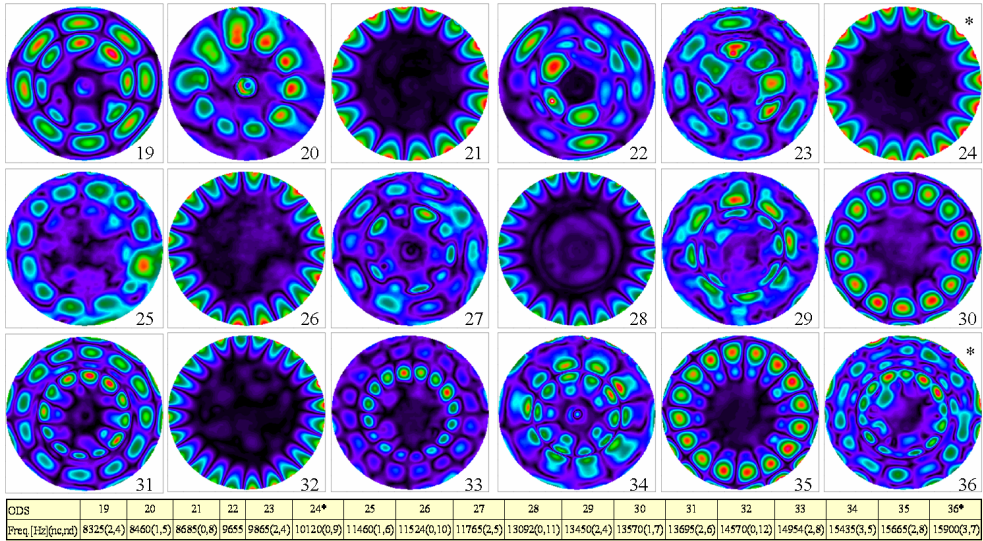


Figure 14: Brake disk dynamics [9] in the frequency range [8325-15900 Hz]

an effective tool for high frequency squeal analysis, where with traditional sensors the shapes might be roughly sampled in space domain, thus losing precious information. High spatial resolution strongly reduces the search for correct measurement points to provide complete and accurate results.

The disc was excited in the out-of-plane direction with a single frequency stepped sine sweeping from 1 to 23 kHz, with variable density around the interesting frequency previously detected by time average techniques [1]. The most interesting results are reported in Figs.(13-15) by complex amplitude; complex phase must be considered as well when judging the shapes. Axisymmetric ODSs, strictly influenced by free disc behaviour, were predominant and more easily detectable. The classical definition (nc, nd), accordingly to the number of nodal circles and diameters can be applied to ODS that reveal strong influence of axisymmetric modes. Other shapes occurred, linked to the underneath structure of the disc and to the influence of different modes. The complex-valued ODS proved to be the right procedure to obtain the best results.

From the coupling point of view, the modes most prone to squeal are those that geometrically match the brake pad eigenshapes, which are very similar to bending and twisting modes of beams or thick plates. It was here possible to assess that many ODS may meet this requirement, leading to the possibility for the underneath modes of the brake disk to be matched with bending, but also twisting and higher cupping modes of the pads. This shows a clear benefit of the high spatial resolution in the measurements provided by ESPI. Rough spatial sampling is avoided in a very wide frequency range, leading to an accurate detection of local shapes also at high frequency and with no need of looking for the best location for measurement dofs. This stands in the perspective of providing useful data to perform experimental modal analysis and model updating for highly detailed models of the squeal phenomenon in the high frequency range, these models being experimental, hybrid or exclusively numerical.

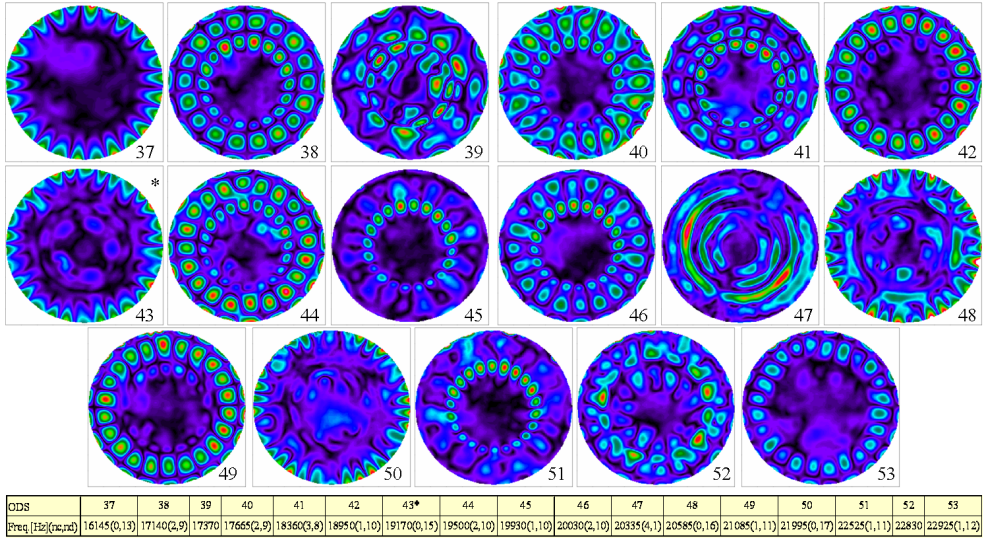


Figure 15: Brake disk dynamics [9] in the frequency range [16145-22925 Hz]

6. APPLICATION: APPROACHING NVH/CAE/EMA [7]

Light structures easily exhibit high modal density with complex patterns, where the shapes might be resolved with a fine grid of experimental dofs and without adding the sensor masses to the structure. Coherent light interferometry techniques, like ESPI/Holography, were initially developed to detect displacement maps in the sub-micron range at a single dynamic event, such as a shock, a single frequency vibration or a static deflection, but they may challenge other experimental techniques also in the vibration field.

The wide use of ESPI technology for NVH topics has still to come: a reason might be the huge amount of data that arises in a wide frequency dataset. Higher resolution in the sensor means increasing the quality of the recordable specklegrams, and thus of the displacement fields, in terms of sensibility and range of measurements. But this increases the size of data, while the number of dofs useful for NVH researches should be linked to the shape of the highest eigenmode considered.

Bridging to other NVH environments

Electronic Holography or ESPI may easily give rise to datasets of about one million dofs. Besides, the use of harmonic excitation and stroboscopic light imposes no restriction on the frequency accuracy, this depending only on the function generator and on the measurement system on-board electronics. But the number of effective degrees of freedom in the dynamic events are only a small portion of the measured responses, so the huge amount of redundant data that follows from ESPI might be reduced, also to cope with present computing powers. A simple approach was implemented, by means of sorting those experimental dofs close to a virtually projected $m \times n$ grid.

The animation of the receptance FRF maps on the object was also implemented frequency by frequency, to better appreciate the complex-valued nature of the data, with its

unevenly distributed phase delay that it may contain. The real part, the imaginary part, amplitude and phase data can thus be animated on the 3D geometry. In Fig.(16) the animated surface is tessellated by triangles from the grid structure; each triangle is coloured by a linear transition of the three colours in the vertex; these latter RGB colours are the expression of the level reached by the corresponding displacement in each global direction, with regard to the range of interest; the red component is taken for x displacement, the green and the blue ones for y and z respectively. The black-dark areas in Fig.(17) are the nodal lines of the ODS, while the bright ones approach the maximum, in the complex-valued amplitude representation. By exploiting fitted external input channel and reduced dof dataset, the FRF data were extracted from the ESPI displacement and contouring maps. The FRFs were packed as frequency uneven functions. FRFs and geometry could then be exported to other NVH frequency domain applications.

Experimental Modal Analysis from ESPI measurements

The specimen under test was an aluminium rectangular plate, fixed by tightened wires to a frame clamped on an air-spring optical table, as in Fig.(6). The excitation was provided in the out-of-plane direction by a shaker.

An extended measurement campaign covered the frequency range 20-1000 Hz, with about 1 Hz of increment per step in the stepped-sine sweeping frequency collection. In the tests carried out, no clear fringe appeared for the in-plane motion, meaning that the amplitude was below the resolution of the instrument, so in the nanometer sub-range. Besides, animating the complex-valued 3D fields as in Fig.(16), it was possible though to appreciate the rough information coming from the shallow prints in in-plane measurements, which underlined the location of the force cell and shaker stinger as a source of heterogeneity for the thin plate.

In the frequency range of interest, it first appeared that the ODSs exhibited by the plate did not present such spatial complexity as to require an extremely refined grid. It seemed that the number of dofs did not introduce any marked enhancement in the z displacement field over the threshold of 20 dofs per plate side. Also, restrictions come from the modal analysis tools, designed for a lower number of dofs, and from the computing power required. The Z -direction receptance FRFs calculated were exported to a commercial modal analysis software together with contouring data. An increased detection ability of the eligible poles in the stability chart followed with the dof number rising. The poles resulted a little more

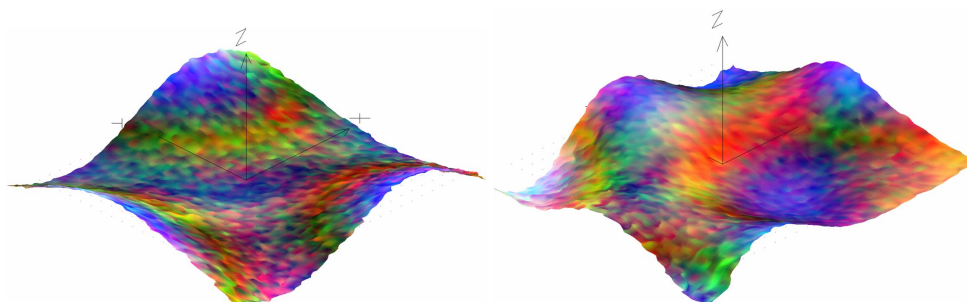


Figure 16: 3D Plate dynamics [7]: a) 308, b) 536 Hz

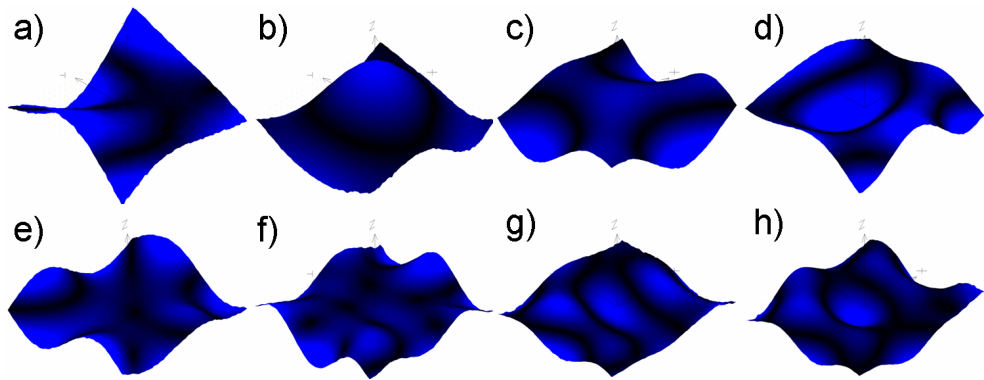


Figure 17: Out-of-plane plate dynamics [7]: a) 115, b) 277, c) 322, d) 394, e) 408, f) 755, g) 857, h) 923 Hz

clear and stable, while it was easier to find meaningful shapes at higher frequencies. Still, the identification was not completely satisfactory and the smoothness of the mode shapes did not reach the one expected from such accurate measured ODSs. The identification algorithm exploited the growing redundancy present in the data better than the simple visual inspection of the ODSs; thus the dof number increase helped also in partially overtaking some problems occurred on the right frequency sampling; however, a data source spatially accurate, but not with an exact frequency information, gave low quality outputs regarding the shapes: therefore dedicated algorithms, image-spatial domain based, might be investigated in future studies.

7. APPLICATION: FATIGUE LIFE ASSESSMENT [10–12]

The dynamic behaviour influences not only the possibility to reach functionality targets, but also the capacity to continuously replicate them and finally the reliability of the components. Integrity or manufacturing quality controls might be insufficient to assure in-service safety: in dynamic loading events a small defect in a highly stressed part of the product can not be tolerated, while a bigger one is not as dangerous in a less fatigued zone. Dynamic analysis becomes strategic for both performance and durability design targets. The ever increasing part of design stage taken by virtual models still needs experimental results to update the simulations and optimize all the parameters, especially when the inherent structural dynamics and boundary constraints raise high modal density with complex displacement patterns and generalized damping. To predict the failure of a complex structure with accuracy, stresses or strains should be known everywhere over the entire structure. This is not commonly achieved with conventional instruments, like the widespread strain gages, which may face some problems, like effective sensor placement, spatial averaging, mass burden and lumped measurements. Multimodal behaviour of the structure under dynamic loading in service increases the strain-gage location assessment uncertainty.

ESPI experimental technique becomes particularly suited for the evaluation of the strain field in dynamic events, when its full-field intrinsic nature easily displays complex-shaped patterns at high frequency, with high-resolution and no need for contacting/mass loading instrumentation that modifies the responses, drastically reducing the failures in locating the sensors. The complex-valued fields enhance the treatment of the experimental results and

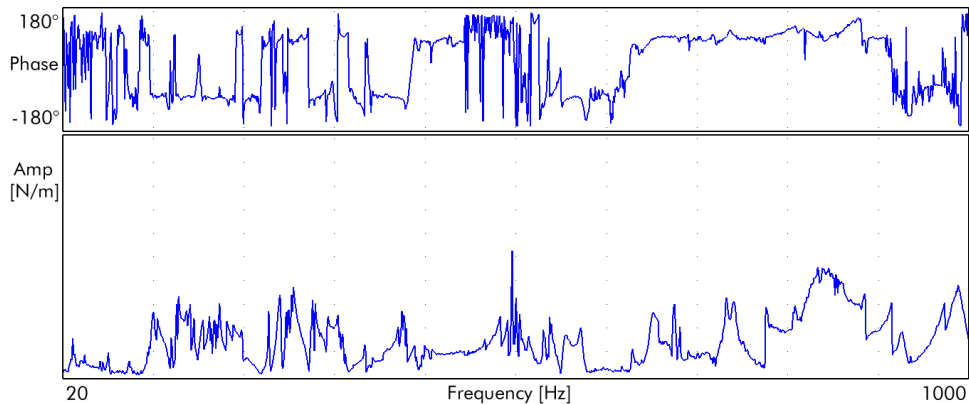


Figure 18: Out-of-plane plate FRF [7]

help avoiding basic physical misunderstandings. Further, by means of receptance FRF-ESPI data, an experiment-based simulation can be run to extend the fatigue life calculation to different broad band spectral and amplitude contents.

Cumulative fatigue life assessment

In long- and giga-cycle- life fatigue (10^6 - 10^9 cycles) low stress-strain is expected. Elastic behaviour of material is assumed and $S - N$ curves are used to relate stress levels to expected number of cycles before crack propagation. In spite of its major shortcomings (absence of loading interaction, sequencing and memory effects, failure with a wide range of the total damage), the basic Palmgren-Miner linear damage rule is still dominantly used in design. In [10,11] the use of full-field measurements was presented as a promising evolution of accepted procedures, to obtain the fatigue life prediction without selecting *a priori* the location of the sensors on the sample and to avoid any simplification into a broad band dynamic behaviour, expected to be populated by complex-shaped patterns due to high modal superposition. The approach proposed by ESDU [13–15], to evaluate fatigue life in low-stress / high-cycles random loading for isotropic metal alloys, was applied and extended over a much higher number of experimental measurement points than in conventional techniques. Besides, many other prediction approaches can proficiently exploit ESPI dynamic measurements.

S - N curves in low & high cycles. For the present applications, the tests on standard isotropic metal samples are obtained by means of reversed bending or acoustic excitation under stress-controlled procedures. The experimental $S - N$ curve that relates the stress range to the sustainable number of cycles can be modelled by means of the exponential Basquin's law:

$$NS_r^b = K_r \quad (5)$$

The parameters (b and K_r) therein are obtained by least-squares fitting of the experimental data in a log-log plot of Number-of-cycles N and Stress-range S_r . The curve of each material can be exploited by comparison of equivalent stress range when multi-axis loading occurs. Depending on the equivalent stress range map and on $S - N$ curve, a distribution of cycles-to-failure is sketched over the sample surface.

Cumulative damage model: a spectral approach. A cumulative damage model and cycle counting procedure is necessary due to the variable amplitude and random loading in broad band excitation. Frequency domain approaches have to be used with ESPI data, exploiting the analogy with rain-flow cycle counting and statistical parameters in time domain. An equivalent range map of stress cycles \mathbf{S}_{eq} , representative of the damage inferred across the surface by the retained dynamics in the whole spectrum, is the result calculated by means of the spectral cumulative damage methods [13–15]. Being $p(\mathbf{S}_r)$ the probability density function (pdf) of \mathbf{S}_r map at frequency ω , adopting the proper $S - N$ curve, the total damage map \mathbf{D} accumulated in N cycles distribution of broad band spectrum loading becomes:

$$\mathbf{D}(N) = \frac{N}{K_r} \mathbf{S}_{eq}^b = \frac{N}{K_r} \int_0^\infty \mathbf{S}_{r,p}^b(\mathbf{S}_r) d\mathbf{S}_r \quad (6)$$

Spectral parameters. The main parameter $\mathbf{m}_k = \int_0^\infty f^k \mathbf{PSD}_{VM}(\omega) d\omega$ is the k -th order moment of the frequency by the power spectral density (PSD) of the VonMises equivalent stress $\mathbf{PSD}_{VM}(\omega)$; $\mathbf{F}_k = (\mathbf{m}_k/\mathbf{m}_0)^{1/k}$ is the k -th characteristic frequency; and $\gamma_k = \mathbf{m}_k / (\mathbf{m}_0 \mathbf{m}_{2k})^{1/2}$ is a whole class of bandwidth parameters to describe the distribution of the stationary and Gaussian processes; $\mathbf{F}_p = (\mathbf{m}_4/\mathbf{m}_2)^{1/2}$ is the expected number of peaks per time unit.

The Dirlik's semi-empirical spectral method was adopted for broad band spectra of stress responses, as suggested in [15], for its proved effectiveness in giving the best prediction of fatigue life. Assuming that failure happens as the total damage reaches the unitary value, the distribution of time-to-failure predictions is obtained as:

$$\mathbf{T}_{failure} = K_r / (\mathbf{F}_p \mathbf{S}_{eq}^b) \quad (7)$$

Strain-stress fields from ESPI measurements

ESPI can deliver a full PSD map of the equivalent stress, to extend the calculation of the spectral method over a fine grid of equivalent sensor locations. Instead of few fatigue life evaluations, there results a fine map of thousands of predictions, referenced to the geometrical location (x, y) in the ESPI-sensor reference; the advantage is strait forward, as the calculations are free from sensor location choices, especially facing multi-modal high complex-shaped displacement patterns measurable at extremely high frequencies, and when exciting interface and real boundaries might need non-trivial models. ESPI high spatial resolution brings also a reduction of sensor area-averaging effect and set-up complexity, as can be found instead with more conventional instruments.

To calculate the PSD of the equivalent stress, from dynamic ESPI displacements, strains need to be evaluated by means of a robust differentiation procedure; a constitutive model must be adopted to calculate the stresses, from which a failure criteria extracts the equivalent stresses frequency by frequency, finally retained in the PSD map requested.

Strains from ESPI displacement fields

ESPI technology delivers measurements independently of the material type. No bulk strains and stresses are obtainable. A plane-stress scenario was adopted in [10, 11] according to the type of measurements obtained from the tested samples.

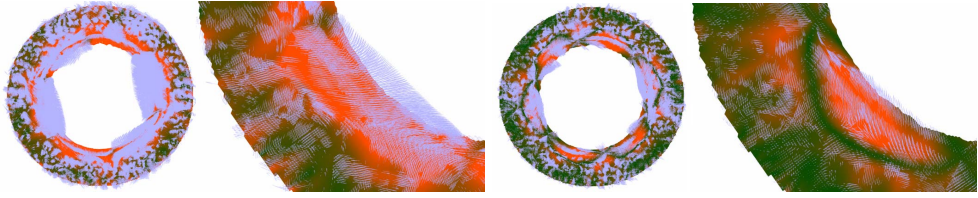


Figure 19: 1st & 2nd principal stresses on a disk at 3707 Hz [10]

Tensor components. At each frequency the complex-valued strain tensor components ε_{ij} can be calculated, according to strain theory, by means of successive derivations of the complex-valued displacement $\mathbf{d}_\omega(x, y)$. In the plane-stress model (engineering notation), the in-plane displacements give their contribution to the tensor components; bending-related in-plane strain results from the out-of-plane displacements.

In a general test-based approach a rough numerical differentiation might greatly amplify the influence of measurement and discretization noise and must therefore be performed with procedures that mitigate the differentiation errors. Digital filtering procedures [16] obtain n -th order derivatives with greatly increased accuracy in respect to simple numerical differentiation. The 0-th order derivative becomes also a powerful tool for smoothing the original data arrays.

Stress fields

To calculate the σ -tensor from ε -tensor a proper constitutive model must be adopted. A low-strain/high-cycle fatigue behaviour was expected from the samples under test, as well as a linear isotropic elastic behaviour of the material. Furthermore, it must be remembered that ESPI technology is limited by the number of resolvable fringes per sensor to low displacement ranges, typical of elastic behaviour; then by means of receptance ESPI-FRF the linearity is imposed on higher displacements ranges.

Constitutive model. A complex-valued linear isotropic constitutive model is outlined, which includes material parameters and transformation between complex-valued strain ε -tensor and stress σ -tensor components. The complex-valued principal strains and stresses can be then easily evaluated, together with their principal directions.

Failure criteria for multi-axis equivalence. At present [17] no general comprehensive theory exists to relate fatigue limits among different stress states; the latter are generally coming from multi-axis loading, while most fatigue tests on materials [18] are run on specimens in a controlled mono-axial loading procedure. An empirical equivalence criteria has thus to be formulated to compare the results. Von Mises (ductile materials) or Tresca (brittle materials) failure criteria are widely recognised to be an acceptable compromise to obtain equivalent stresses. Equivalent complex-valued stresses are evaluated in each (x, y) location of ESPI maps, from the principal components maps. The full-field 3D intrinsic nature of ESPI allows to discard the calculation of notch stress concentration factor: a complex-valued map of the principal stresses and directions reduces the uncertainty in the stress distribution, especially in multi-axis loading sequences.

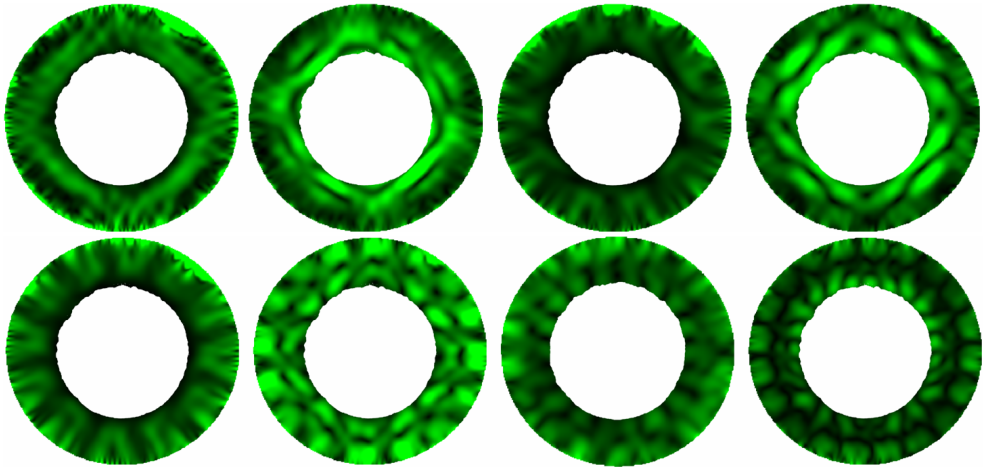


Figure 20: VonMises stresses on a disk at 1100, 3707, 7260, 8325, 8685, 15665, 17650, 20030 Hz [10, 12]

Application: fatigue life assessment of a brake disk and a plate

From the viewpoint of computing requirements, at this stage of software development, the procedure is demanding: four estimations run in parallel took about 24 hours on a dual Intel-Xeon-dual-core-processor workstation, with dynamic allocation of about 1 Gbyte of memory for each process to manage all the data; more work has to be done on this side to increase the overall computational performance.

The author reported in [9] the results of an intensive measurement campaign to accurately describe the behaviour of the disk in as many high spatial resolution ODSs as possible at the frequency lines of a broad spectrum; in [10] the fatigue life assessment of the same brake disk was discussed and here the results are recalled. The author worked on plate dynamic behaviour in [7], where broad band displacement measurements were joined with excitation channel sampling to calculate the maps of receptance FRFs. The obtained results were used in [11] to assess the expected time-to-failure of the plate: Dantec Ettemeyer Q500 VibroESPI was used for all the measurements.

Material parameters. The following parameters were chosen for the cast iron constitutive model of the brake disk: $E= 1.04e+11$ Pa; $\nu= 0.26$. The material fatigue specific parameters were $b= 4.762$ and $K_r= 7.9085e+42$ Pa^b. A grid of 256x256 dofs was kept for further evaluations.

The constitutive model and $S - N$ curve parameters of the plate were taken from the Aluminium 7075T6 properties [15]: $E= 7.00e+10$ Pa, $\nu= 0.33$, $b= 4.810$, $K_r= 4.4200e+43$ Pa^b. To be able to calculate the receptance FRF, force signal was sampled at the shaker-plate interface. The plate data were decimated to 40804 dofs and the analyses regarded the inner 200x200 mm area.

Displacement, strain & stress fields. In both set-ups the strain calculation was function of the bending-related displacements only. A robust numerical derivation procedure

Table 1: Displacement and receptance FRF ranges on the plate [11]

Displacement [μm]				
Freq. band [Hz]	Min <i>Re</i>	Max <i>Re</i>	Min <i>Im</i>	Max <i>Im</i>
31-500	-0.361	+0.472	-0.511	+0.589
256-756	-0.406	+0.472	-0.418	+0.505
500-1013	-0.406	+0.461	-0.509	+0.624
31-1013	-0.406	+0.472	-0.511	+0.551
Receptance FRF [$\mu\text{m}/\text{N}$]				
Freq. band [Hz]	Min <i>Re</i>	Max <i>Re</i>	Min <i>Im</i>	Max <i>Im</i>
31-500	-659.0	+613.3	-545.1	+813.4
256-756	-188.4	+160.6	-156.8	+160.6
500-1013	-100.0	+106.6	-32.3	+33.8
31-1013	-659.0	+613.3	-479.4	+429.2

was used to calculate complex-valued strains. Principal complex-valued stresses are pictured in Fig.(19) with oriented arrows over their amplitude colour maps. The Von Mises equivalent complex-amplitude stress distribution is shown in Fig.(20) at some frequencies; in each image black is the lowest amplitude, while bright green is the maximum one. The complex-valued dynamic measurement approach was able to catch 41 rotating/travelling waves on the disk, which being an axis-symmetric structure is prone to show many coincident and orthogonal modeshapes. No force measurements were taken at the shaker-disk interface and the range of displacement amplitudes did not exceed $2 \mu\text{m}$ in the whole frequency band.

In Tab.(1) the ranges obtained in the direct measurements and in calculated receptance FRF maps on the plate are listed; each spanned frequency interval contained 75 unevenly sorted data fields. The implemented procedure, used to filter and derive the ESPI maps, gave in [11] promising results on strains and equivalent stresses. The uniformity of Von Mises stress in some ODSs as in Fig.(21) appeared to be function of the high complexity in the value-nature of data fields, as dispersion on the complex plane: almost all the surface of the plate was interested by comparable complex-amplitude tension levels.

The calculation of the receptance FRF maps becomes basic in the experimental modelling of the linear structural responses: new fatigue life predictions can be calculated by varying the loading condition by means of the complex-valued spectrum of the input.

Cumulative fatigue calculation from dynamic ESPI measurements. Dirlik semi-empirical spectral method can be thus followed to obtain the map of time-to-failure for each sample and loading spectrum.

Table (2) and the maps in Fig.(22) illustrate the results obtained in [10] with different ODS sets taken on the disk. In all the test cases the failure was diagnosed on the outer border, but it took different locations on this latter. For the estimation of the fatigue life of the braking area the high frequency ODSs are relevant, as revealed by the time-to-failure distribution of Fig.(22) where the brighter colour indicates higher fatigue damage; in particular, the underneath structure of the ribs, delimiting the ventilating channels, appears sensitive to fatigue problems. The ODS *set4* has a time-to-failure of a magnitude lower than when considering the *set2*, and even two orders compared to the *set1*; high frequency structural

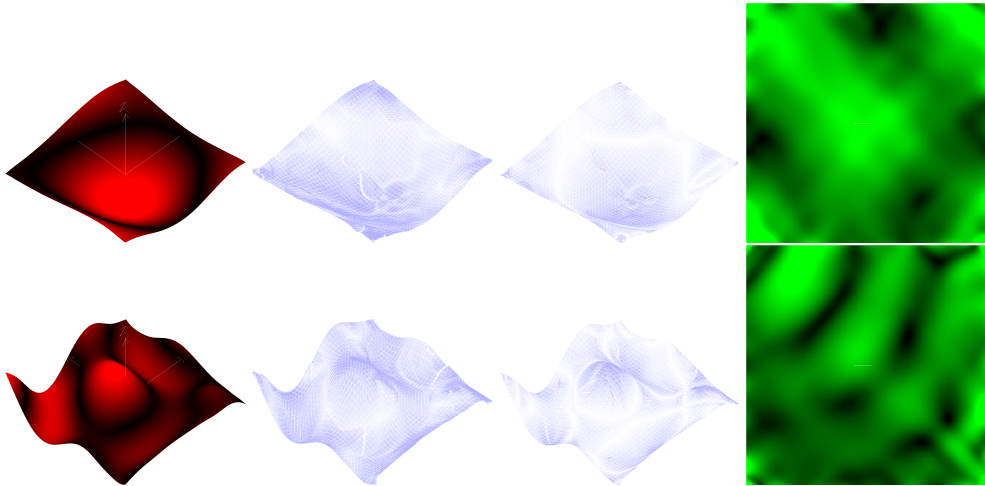


Figure 21: ODS, 1st & 2nd Principal and Von Mises stresses at 272 and 948 Hz [11, 12]

dynamics confirms to be of uttermost importance in developing fatigue life assessments.

The estimated life distributions of the plate in [11] are summarized on Tab.(3) and Figs.(23,24); acquired displacement maps and calculated receptance FRF maps were used. The receptance FRFs permit the simulation of the component dynamics as subjected to an unitary amplitude input force with white noise spectrum. The structural dynamics retained in the calculations seems again to play a relevant role in determining the distributions of critical locations inside the area. High levels of Von Mises stresses were reached by means of receptance FRF at some lower frequency lines, drastically reducing the expected time-to-failure to instant crack propagation; this was predictable also looking at the ranges in Tab.(1) and at the input force ranges.

The calculations done with receptance FRFs, since they allow the simulation, within the linearity assumption, of much higher strains than those measurable by means of ESPI dynamic measurements, might be processed in the frame of total strain energy related approaches. Not only the white noise unitary force input can be simulated by means of the receptance FRF, but also any other input spectra.

Fatigue estimation notes

All the tests were run in the frame of non-contact full-field optical techniques: the estimation of equivalent stresses in the wide frequency band process was obtained without identifying *a priori* any best sensor location. This latter aspect might become of uttermost importance in dealing with detailed multi-modal behaviour with complex-shaped patterns at high frequencies. The time-to-failure distribution of the component under test was evaluated without the need of any analytical/numerical structural model, but only robust numerical differentiation, material constitutive laws and Dirlik's spectral cumulative damage model. ESPI-based procedure has clearly revealed influences of the structural dynamics on the time-to-failure distribution as different value ranges and distribution morphology. Further, the evaluation of the experimental receptance FRF maps becomes a promising approach to simulate fatigue

Table 2: Assessed time-to-failure distributions on the disk in hours [10]

Set	ODS #	Band[Hz]	Min	[dof]	Max	[dof]	Mean	std. dev.
set1	200	670-3162	5.95E+04	6713	2.97E+19	40601	1.18E+17	6.93E+17
set2	20	1100-8460	2.76E+03	5568	3.45E+146	34689	8.47E+10	2.28E+12
set3	20	15435-22925	1.58E+05	5825	1.15E+16	34164	4.04E+13	3.72E+14
set4	53	1100-22925	3.70E+02	5825	1.04E+14	34689	2.31E+10	8.06E+11

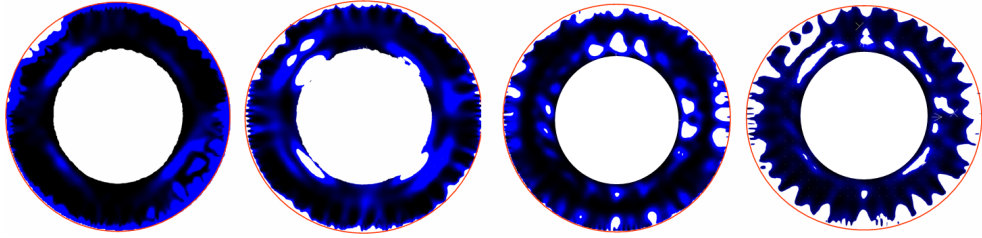


Figure 22: Fatigue location on the brake disk in the spectra sets: *set1*, *set2*, *set3*, *set4* of Tab.(2) [10]

life as function of the varying input spectra without the aid of structural or virtual models. Certainly, verification of life estimates should be considered an important activity to confirm the calculations and the relevance played by a proper modelling of the higher frequency structural dynamics and multi-axis loading patterns.

8. CONCLUSIONS AND FUTURE DEVELOPMENTS

The review showed the great potential that ESPI, as optical full-field technique, might have in many different branches of structural dynamics. Still full-field measurements, not only ESPI-based, deserve more research effort to be exploited at their best: future attention will be paid to improve the computational efficiency of the processing environment in the perspective of modal analysis, modal strains recovery and vibroacoustic modelling; other full-field measurement techniques will be investigated, starting from the basics of signal processing.

Acknowledgements

The measurements here processed were acquired during a European Commission - Research Directorate General - Marie Curie Post Doctoral Research Host Fellowship at Dantec Ettmeyer GmbH, Ulm, Germany, under the project HPMT-CT-1999-00029 "Speckle Interferometry for industrial needs", 1/4/2004-30/6/2005. This work is intended to be part of the spreading of the results thereof. Both European Commission and Dantec Dynamics (former Dantec Ettmeyer) are acknowledged.

REFERENCES

- [1] Doval, A. F., 2000. "A systematic approach to tv holography". *Measurement Science and Technology*, **11**(1), pp. R1–R36. IOP Pub Ltd.

Table 3: Assessed time-to-failure distributions on the plate in hours [11]

Displacement						
Band[Hz]	Min	[dof]	Max	[dof]	Mean	std. dev.
31-500	3.21e+02	56525	6.49e+08	30862	7.78e+07	9.51e+07
256-756	6.52e+01	6476	2.03e+08	39092	1.99e+07	3.02e+07
500-1013	4.29e+00	56726	7.24e+06	25587	8.63e+05	9.51e+05
31-1031	1.71e-01	56525	6.59e+05	25336	8.92e+04	9.37e+04

Receptance FRF						
Band[Hz]	Min	[dof]	Max	[dof]	Mean	std. dev.
31-500	6.41e-11	6476	9.53e-04	17395	1.17e-04	1.40e-04
256-756	2.67e-07	6476	7.33e-01	30626	2.96e-02	7.05e-02
500-1013	1.65e-07	56726	1.44e+00	41345	7.29e-02	9.91e-02
31-1031	7.03e-11	6275	1.52e-04	12614	2.23e-05	2.62e-05

- [2] Rastogi, P. K., 1997. *Optical Measurement Techniques and Applications*. Artech House, Inc., Nordwood, MA 02062, U.S.A.
- [3] Maas, A. A. M., 1991. "Phase shifting speckle interferometry". PhD thesis, Delft University Press.
- [4] Vrooman, H. A., 1991. "Quantitative analysis of interferograms". PhD thesis, Delft University Press.
- [5] Kreis, T., 2004. *Handbook of Holographic Interferometry*. Wiley-VCH, Berlin, Germany.
- [6] Zandarini, A., 2005. "Damage location assessment in a composite panel by means of electronic speckle pattern interferometry". In Proceedings of the IDETC/CIE ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Long Beach, California, USA, September 24-28, ASME. Paper DETC2005-84631.
- [7] Zandarini, A., 2007. "Full field espi measurements on a plate: challenging experimental modal analysis". In Proceedings of the XXV IMAC, Orlando (FL) USA, Feb 19-22,, SEM. Paper s34p04.
- [8] Ewins, D. J., 2000. *Modal Testing - theory, practice and application*, 2nd ed. Research Studies Press Ltd., Baldock, Hertfordshire, England.
- [9] Zandarini, A., 2005. "Dynamic behaviour characterization of a brake disc by means of electronic speckle pattern interferometry measurements". In Proceedings of the IDETC/CIE ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, Long Beach, California, USA, September 24-28, ASME. Paper DETC2005-84630.
- [10] Zandarini, A., 2008. "Fatigue life assessment by means of full field espi vibration measurements". In Proceedings of the ISMA2008 Conference, September 15-17, Leuven (Belgium), P. Sas, ed., KUL, pp. 1-15. Paper 326.
- [11] Zandarini, A., 2008. "Full field espi vibration measurements to predict fatigue behaviour". In Proceedings of the IMECE2008 ASME International Mechanical Engineering Congress and Exposition, October 31- November 6, Boston (MA) USA, ASME, pp. 1-10. Paper IMECE2008-68727.
- [12] Zandarini, A., 2008. "Esp measurements in structural dynamics: fatigue life assessment". In Proceedings of the Dantec Dynamics 15th International Conference and User Meeting, September 22-23, Ulm, Germany, Dantec Dynamics, pp. 1-20. ISBN: 978-3-00-025353-9.

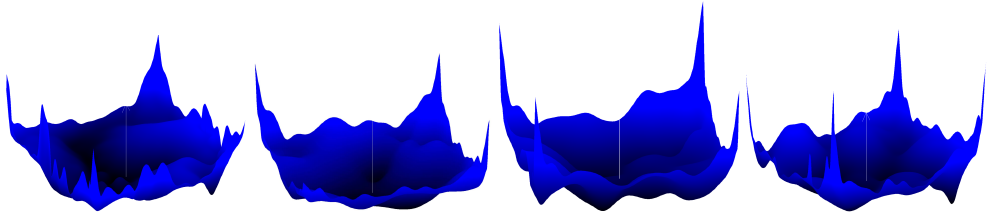


Figure 23: Fatigue damage distributions from ESPI measurements (31-500, 256-756, 500-1013 and 31-1013 Hz) [11, 12]

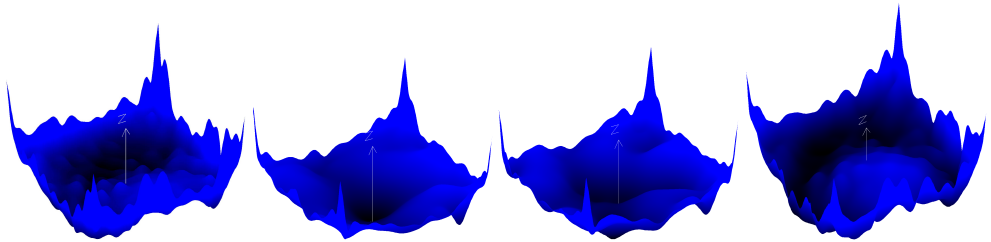


Figure 24: Fatigue damage distributions from receptance ESPI-FRF (31-500, 256-756, 500-1013 and 31-1013 Hz) [11, 12]

- [13] ESDU, 2006. “Fatigue damage and life under random loading 06009”. In *Estimation of fatigue life of structures subjected to random loading*, Vol. 9 of *Vibration and Acoustic Fatigue Series*. ESDU International Plc, London, May.
- [14] ESDU, 2006. “Cycle counting methods for the estimation of fatigue life 06010”. In *Estimation of fatigue life of structures subjected to random loading*, Vol. 9 of *Vibration and Acoustic Fatigue Series*. ESDU International Plc, London, August.
- [15] ESDU, 2006. “Spectral methods for the estimation of fatigue life 06011”. In *Estimation of fatigue life of structures subjected to random loading*, Vol. 9 of *Vibration and Acoustic Fatigue Series*. ESDU International Plc, London, November.
- [16] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992. *Numerical Recipes in C - The Art of Scientific Computing*, 2nd ed. Cambridge University Press, England.
- [17] McDowell, D. L., 1996. “Multiaxial fatigue strength”. In *Fatigue and Fracture*, Vol. 19 of *ASM Handbook*. ASM International Handbook Committee.
- [18] George, T. J., Seidt, J., Shen, M. H. H., Nicholas, T., and Cross, C. J., 2004. “Development of a novel vibration-based fatigue testing methodology”. *Int. J. Fatigue*, **26**, pp. 474–486. Elsevier Science Ltd.

ANALISI A COLLASSO DI GIUNZIONI INCOLLATE: MODELLAZIONE EFFICIENTE AGLI ELEMENTI FINITI E CONVALIDA SPERIMENTALE

Eugenio Dragoni

*Dipartimento di Scienze e Metodi dell'Ingegneria
Università degli Studi di Modena e Reggio Emilia
E-mail: eugenio.dragoni@unimore.it*

Davide Castagnetti

*Dipartimento di Scienze e Metodi dell'Ingegneria
Università degli Studi di Modena e Reggio Emilia
E-mail: davide.castagnetti@unimore.it*

Andrea Spaggiari

*Dipartimento di Scienze e Metodi dell'Ingegneria
Università degli Studi di Modena e Reggio Emilia
E-mail: andrea.spaggiari@unimore.it*

Abstract. *Il lavoro descrive una procedura innovativa ed efficiente basata su tecniche agli elementi finiti standard per l'analisi a collasso di strutture incollate. Il lavoro verifica l'accuratezza del metodo proposto nella previsione della risposta post-elastica di giunti incollati, applicando sia un criterio di cedimento alle tensioni regolarizzate che un criterio di cedimento della zona coesiva, rispettivamente a modelli piani e tridimensionali. La verifica del modello si svolge mediante il confronto con le curve forza spostamento ottenute da una campagna sperimentale sviluppata ad-hoc su giunzioni T-peel. L'accuratezza del modello appare molto buona rispetto ai risultati sperimentali, sia in termini di forza massima che di comportamento post-elastico. Sia il criterio di cedimento alle tensioni regolarizzate che quello della zona coesiva appaiono appropriati per il modello computazionale. Il tempo di CPU impiegato per l'analisi, inoltre, è minimo e questo convalida l'efficienza della procedura per l'analisi di strutture molto complesse.*

Keywords: *analisi a collasso, elementi finiti, giunti T-peel, prove sperimentali, tensioni regolarizzate, zona coesiva.*

1. INTRODUZIONE

Il lavoro riguarda la verifica di un modello ridotto agli elementi finiti (EF), precedentemente verificato dagli autori in campo elastico, per l'analisi a collasso di giunzioni strutturali incollate. L'obiettivo del lavoro è duplice: in primo luogo valutare l'accuratezza del modello computazionale nella previsione della risposta post-elastica; in secondo luogo verificare l'applicabilità di un criterio di cedimento alle tensioni regolarizzate per descrivere l'evoluzione del cedimento nell'adesivo fino al collasso del giunto.

La motivazione della ricerca risiede nel fatto che l'applicazione delle giunzioni strutturali incollate dipende dalla capacità di prevedere la loro resistenza meccanica per mezzo di metodi di calcolo semplici, veloci e accurati.

In letteratura si ritrovano molti metodi agli elementi finiti per l'analisi delle giunzioni incollate [1-4]. Tutti questi metodi sono basati su elementi speciali per descrivere lo strato adesivo o la zona di sovrapposizione. In lavori recenti i metodi più comunemente impiegati adottano approcci alla meccanica della frattura. I principali svantaggi di questi metodi risiedono nel fatto che gli elementi speciali da essi impiegati sono difficili da implementare nei software agli elementi finiti commerciali impiegati nell'ambito industriale e il loro uso è confinato ad applicazioni di ricerca. Inoltre, i criteri basati sulla meccanica della frattura richiedono dati che sono difficili da ottenere sperimentalmente.

Per superare queste limitazioni, questo lavoro propone un metodo computazionale semplificato. Basato su strumenti di modellazione standard ed elementi finiti comuni, implementati nella maggior parte dei software di calcolo commerciali. Un precedente lavoro pubblicato dagli autori [5] ha dimostrato l'efficienza e l'accuratezza del modello ridotto nel calcolare la distribuzione delle tensioni elastiche lungo il piano medio dello strato adesivo per parecchie geometrie 2D e 3D. Successivamente, gli autori hanno esteso il metodo in campo post-elastico [6] adottando il semplice criterio di cedimento alle tensioni regolarizzate proposto in [7] ed ottenendo risultati incoraggianti.

Questo lavoro estende le precedenti analisi, applicando il metodo ridotto ed il criterio di cedimento esaminato, ad un ambito di configurazioni di giunzione più ampio, caricate molto oltre il limite elastico. Il metodo è applicato sia come modello piano che come modello 3D. In quest'ultimo caso si è implementato un criterio di cedimento secondo l'approccio della zona coesiva per aumentare l'accuratezza del modello. Elemento di confronto per le analisi computazionali è rappresentato dalle curve forza spostamento ottenute da una campagna sperimentale sistematica su giunzioni T-peel caricate a collasso.

L'originalità del lavoro consiste nella semplicità degli strumenti computazionali proposti, che sono basati su opzioni standard di modellazione disponibili in ogni pacchetto di calcolo agli elementi finiti commerciale. Questo rende il metodo uno strumento generale e facile da impiegare, caratterizzato da una sostanziale riduzione del costo computazionale dovuta alla minimizzazione dei gradi di libertà del modello. Un secondo vantaggio del lavoro è la convalida, nel caso di modello piano, del criterio di cedimento alle tensioni regolarizzate per descrivere il progressivo danneggiamento dell'adesivo nel dominio post-elastico.

2. MATERIALI E METODI

Il lavoro è diviso in due fasi: prove sperimentali ed analisi computazionali, queste ultime condotte su geometrie sia piane che tridimensionali. E' stata considerata una configurazione di giunzione di tipo T-peel, (Fig. 1a), in virtù del suo collasso progressivo quando è caricata oltre il limite elastico. Il collasso graduale di questa giunzione permette un semplice controllo della curva forza-spostamento fino alla completa rottura, sia nel caso sperimentale che nelle analisi computazionali.

Campagna sperimentale

La campagna sperimentale è stata pianificata secondo i criteri della progettazione degli esperimenti (DOE). E' stato sviluppato un piano fattoriale completo considerando due livelli per ciascuna delle seguenti tre variabili (vedi Tabella 1 e Fig. 1a): spessore degli

aderendi, t , lunghezza dell'incollaggio, l , e materiale degli aderendi. Combinando in tutti i modi possibili i valori di queste variabili, si sono ottenute otto configurazioni di giunzione. Si sono svolte cinque ripetizioni per ciascuna configurazione di giunzione per un totale di 40 giunti. I materiali usati per realizzare gli aderendi sono alluminio e acciaio e l'adesivo impiegato un epossidico bi-componente ad alta resistenza (Henkel 9466 [8]). La Tabella 1 raccoglie le proprietà elastiche degli aderendi e dell'adesivo mentre il loro comportamento post-elastico è descritto dalle curve di Fig. 2a e 2b. Le proprietà post-elastiche degli aderendi sono state ottenute attraverso test sperimentali di trazione. La larghezza dei giunti è di 25 mm e lo spessore dello strato adesivo s_a è stato mantenuto costante a 0.1mm. I test sperimentali sono stati svolti quasi staticamente ad una velocità costante della traversa di 1 mm/min fino al collasso completo del giunto.

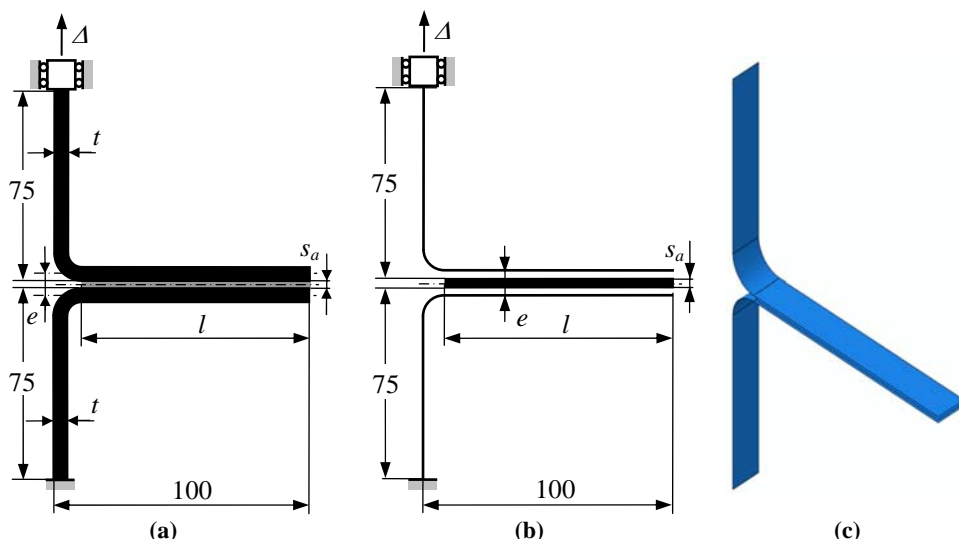


Figura 1. – Schema del giunto T-peel (a) e del modello computazionale 2D (b) e 3D (c)

Tabella 1 – Proprietà elastiche aderendi ed adesivo

Geometria				
t (mm)	2		3	
s_a	0.1		0.1	
l (mm)	88		80	
Materiali				
Aderendi			Adesivo	
	Alluminio	Acciaio	Henkel Loctite 9466	
Modulo di Young (MPa)	69.000	206.000	1718	
Modulo di Poisson	0.3	0.3	0.3	
Tensione di snervamento (MPa)	139	295 (2mm) – 500 (3mm)	26.5 (alluminio)	128.6 (acciaio)

Criterio di cedimento per l'adesivo

Per l'adesivo sono stati impiegati due distinti criteri di cedimento. Per il modello computazionale 2D è stato impiegato un criterio di cedimento alle tensioni regolarizzate. Questo approccio è ispirato ai modelli strutturali disponibili in letteratura per l'analisi strutturale dello strato adesivo [9]. Anche Goglio et al. [7] hanno proposto un modello strutturale per la determinazione della resistenza statica di giunti strutturali incollati. Qui il metodo è stato esteso al fine di determinare l'avanzamento del cedimento nello strato adesivo fino al completo collasso del giunto. Il cedimento dell'adesivo si verifica quando la tensione di picco secondo Von Mises, sulla mezzeria dello strato adesivo, raggiunge il valore della tensione di snervamento dell'adesivo. Questa tensione di Von Mises è stata calcolata in corrispondenza della forza sperimentale massima, mediante un'analisi computazionale.

Nel caso del modello computazionale 3D (applicato solo alle giunzioni con aderendi di acciaio) si è impiegato un legame basato sull'approccio della zona coesiva. Tale legame è descritto nel diagramma di Fig. 2c, in termini di relazione tra tensione e spostamento δ . Il criterio combina al suo interno sia l'approccio alle tensioni (nel tratto elastico) che un approccio energetico (nella fase post-elastica). I parametri che governano il cedimento sono pertanto la tensione critica massima, il modulo elastico e l'energia di frattura dell'adesivo. In questo caso si è considerato solo il modo I e, quindi, si è ricavata la tensione normale massima a peel di 101.0 MPa attraverso analisi elastiche preliminari simili a quelle considerate nel caso bidimensionale.

L'energia di frattura considerata è, invece, $G_I = 690 \text{ (J/m}^2\text{)}$ come riportato in [10] sulla base di una caratterizzazione sperimentale.

L'evoluzione del danno nell'adesivo è stata descritta mediante una legge esponenziale con degrado massimo delle proprietà nel tratto post-elastico. La legge esponenziale è stata scelta perché comporta un degrado brusco delle proprietà meccaniche dell'adesivo al superamento del limite elastico e quindi rappresenta un comportamento fragile.

Analisi computazionale

La Figura 1b-c rappresenta schematicamente il modello computazionale in esame. Il modello computazionale è stato sviluppato innanzitutto in forma piana (Figura 1b) e, per le giunzioni con aderendi di acciaio, anche come modello 3D (Figura 1c). Si è scelto di non sviluppare il modello 3D per le giunzioni con aderendi in alluminio a causa della elevata non linearità di risposta che questi hanno presentato sperimentalmente (vedi Risultati e Discussione) che ne rende difficile e poco significativa la previsione computazionale. Il modello bidimensionale (Figura 1b) riproduce gli aderendi per mezzo di elementi strutturali trave 2D, giacenti sul piano medio degli aderendi reali, mentre l'adesivo è modellato mediante una singola riga di elementi continui 2D in tensione piana. La connettività tra la mesh degli aderendi e quella dell'adesivo è ottenuta impiegando vincoli cinematici interni che rendono uguali i gradi di libertà corrispondenti delle parti vincolate. Sia gli aderendi sia l'adesivo, inoltre, sono stati modellati per mezzo di elementi quadratici ad integrazione completa, aventi forma quadrata con lunghezza del lato pari ad un quarto della distanza tra i piani medi degli aderendi ($e/4$).

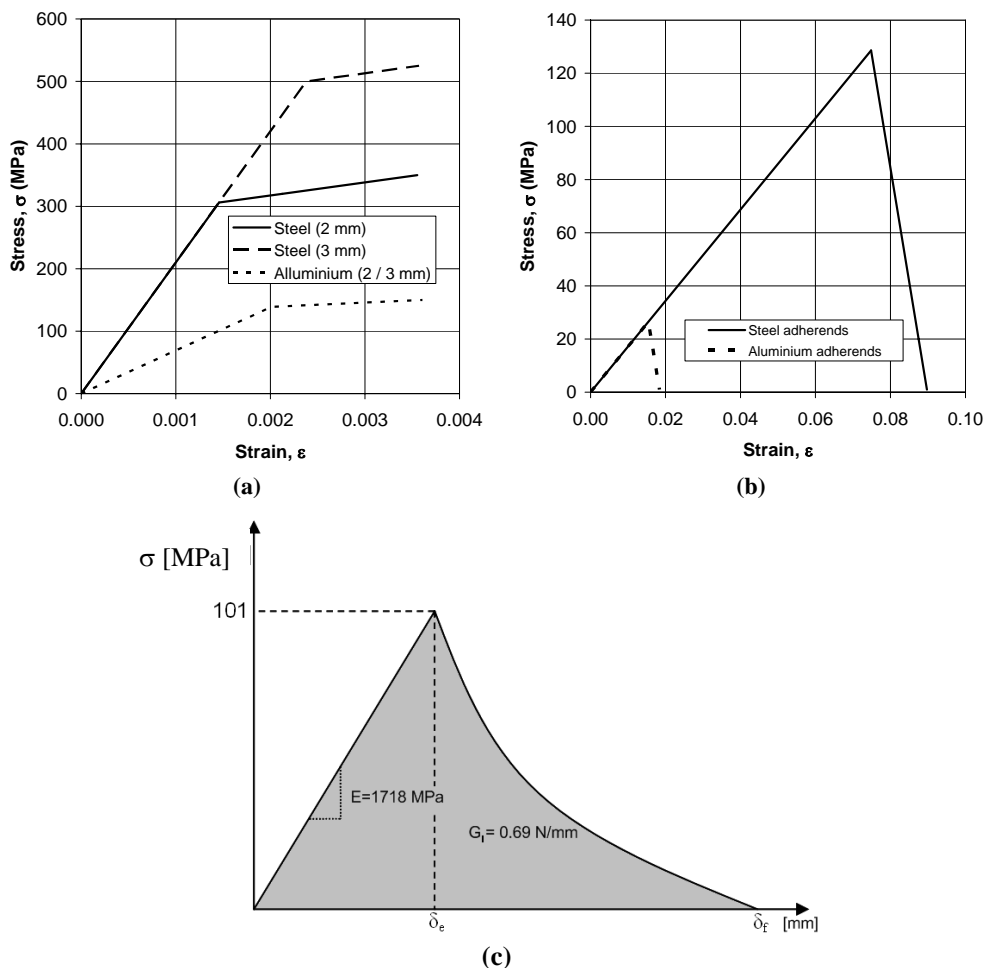


Figura 2 – Curve tensione deformazione adottate nell’analisi agli elementi finite per gli aderendi (a) e per l’adesivo (b) e modello di zona coesiva adottato

I modelli computazionali sono stati sviluppati per tutte le configurazioni esaminate sperimentalmente e sono stati implementati mediante il software agli elementi finiti ABAQUS 6.7 [11]. Il modello tridimensionale (Figura 1c) si differenzia da quello piano innanzitutto in relazione al tipo di elementi impiegati. Gli aderendi sono descritti mediante elementi semistrutturali guscio 3D giacenti sul piano medio degli aderendi ed aventi dimensione del lato pari ad $e/2$. L’adesivo è modellato mediante un singolo strato di elementi coesivi, solidi e continui, aventi lato di dimensione pari ad $e/4$.

Le analisi agli elementi finiti sono state svolte in tre passi, i primi due riguardanti il modello piano ed il terzo relativo al modello tridimensionale. Nel primo passo, avente l’obiettivo di valutare per ciascun giunto la tensione di cedimento dell’adesivo secondo Von Mises, si è descritto il comportamento degli aderendi mediante un legame elastoplastico (vedi Fig. 2a), mentre il comportamento dell’adesivo è stato descritto come

elastico. Il modello riproduce l'intera geometria al fine di ottenere i punti di integrazione esattamente sulla superficie media dello strato adesivo. La presenza di punti di integrazione aggiuntivi, nello strato adesivo, fuori dalla mezzeria dell'adesivo non compromette l'accuratezza del criterio di cedimento perché il valore di tensione risulta all'incirca costante nello spessore di adesivo. Come rappresentato in Figura 1b si è vincolata l'estremità dell'aderendo inferiore con un incastro e quella dell'aderendo superiore con un carrello. Il carico, applicato all'estremità dell'aderendo superiore, è pari alla forza massima misurata sperimentalmente che provoca l'inizio del cedimento del giunto. Questo permette di valutare e confrontare le tensioni di cedimento di tutte le configurazioni di giunzione, permettendo una verifica per il criterio di cedimento adottato.

Lo scopo del secondo e terzo passo dell'analisi computazionale è stato, invece, di ottenere la curva forza spostamento fino al collasso completo, permettendo un confronto diretto con i risultati sperimentali. Nel secondo passo per gli aderendi è stato mantenuto il legame costitutivo elasto-plastico (Fig. 2a) e l'adesivo è stato descritto, invece, mediante i modelli elastico fragili di Fig. 2b. I valori della tensione di cedimento dell'adesivo di Fig. 2b sono stati ottenuti dal primo passo dell'analisi, sia per gli aderendi di acciaio che per quelli di alluminio. Si è applicato uno spostamento Δ al giunto, come mostrato in Fig. 1a, fino alla completa separazione degli aderendi.

Nel terzo passo si è implementato il modello tridimensionale applicando al giunto il medesimo spostamento Δ imposto nel passo precedente, fino alla completa separazione degli aderendi. Per gli aderendi è stato mantenuto il comportamento elasto-plastico descritto in Fig. 2a. L'adesivo, invece, è stato descritto mediante l'approccio di zona coesiva descritto in Fig. 2c, che si rende necessario per gli specifici elementi coesivi impiegati nella sua modellazione.

Tutti i modelli sono stati risolti mediante un processore Intel Core Duo Mobile T7200.

3. RISULTATI

La Tabella 2 raccoglie i risultati delle prove sperimentali e delle analisi computazionali. La prima riga presenta le forze di picco ottenute dalle prove sperimentali di trazione sulle giunzioni T-peel. La seconda riga presenta le tensioni di cedimento calcolate nel primo passo dell'analisi computazionale per le specifiche configurazioni di giunzioni e la terza riga mostra i valori medi di queste tensioni di cedimento per gli aderendi in alluminio ed acciaio rispettivamente. La quarta riga riporta i valori di forza massima ottenuti dal modello 2D agli elementi finiti proposto al secondo passo dell'analisi computazionale per tutte le configurazioni di giunzione considerate. La quinta riga di Tabella 2 presenta un confronto in termini di errore relativo tra il valore di picco sperimentale e numerico 2D. La sesta riga, analoga alla quarta, riporta i valori di forza massima ottenuti dal modello 3D agli elementi finiti, proposto al terzo passo dell'analisi computazionale. Nell'ultima riga, infine, si riportano gli errori relativi tra la previsione della forza del modello computazionale 3D e quella del modello sperimentale.

La Figura 3a-b mostra le curve forza-spostamento ottenute dalle prove sperimentali (linee continue sottili), dal modello computazionale 2D (linea spessa tratteggiata) e dal modello computazionale 3D (linea spessa continua). Per brevità, la Fig. 3 mostra le curve per sole quattro specifiche configurazioni di giunto: un giunto con aderendi di alluminio, spessore 2 mm, con una lunghezza di incollaggio di 88 mm (Fig. 3 a); un giunto con aderendi di alluminio, spessore 2 mm, con una lunghezza di incollaggio di 80 mm (Fig. 3 b); un giunto con aderendi di acciaio, spessore 2 mm, con una lunghezza di incollaggio di

88 mm (Fig. 3c); un giunto con aderendi di acciaio, spessore 3 mm, con una lunghezza di incollaggio di 88 mm (Fig. 3d). Il tempo medio di calcolo è stato di 17.6 s per il modello bidimensionale del secondo passo di analisi e di 450 s per il modello tridimensionale.

4. DISCUSSIONE

Si osserva innanzitutto che i giunti con aderendi di alluminio hanno manifestato una notevole dispersione della risposta sperimentale, sia in relazione alle forze di picco (Tabella 2) che in relazione alla risposta nel tratto post-elastico (Fig. 3a). Nel tratto post-elastico si osserva un andamento insolito se confrontato con la risposta comune di un giunto T-peel. Questo comportamento può essere probabilmente attribuito all'alto grado di non linearità plastica che caratterizza gli aderendi di alluminio durante il cedimento del giunto. La dispersione nella risposta dei giunti con aderendi di acciaio, al contrario, è molto contenuta (Tabella 2) e anche il comportamento post-elastico è simile alla comune risposta del giunto T-peel (Fig. 3b).

Lo stesso livello di dispersione si ritrova nelle tensioni di cedimento dell'adesivo (Tabella 2), calcolate nel primo passo dell'analisi agli elementi finiti per l'intera serie di configurazioni di giunzione. Da un'analisi statistica delle tensioni di cedimento si è osservato che nessuna delle variabili considerate (t , l), ad eccezione del materiale degli aderendi, influisce significativamente sulla tensione di cedimento dell'adesivo. Questo ha permesso l'impiego di un'unica tensione media di cedimento per lo stesso materiale degli aderendi (Tabella 2). E' interessante osservare che la tensione media di cedimento ottenuta per i giunti con aderendi in alluminio è circa 5 volte inferiore di quella ottenuta per i giunti con aderendi di acciaio. Questo può essere probabilmente attribuito alle scarse proprietà adesive dell'alluminio rispetto all'acciaio.

Tabella 2 – Risultati delle prove sperimentali e delle analisi computazionali

		Alluminio			Acciaio	
		1	t = 2 mm	t = 3 mm	t = 2 mm	t = 3 mm
1	Forza di picco sperimentale F_{exp} (N)	80	99.7 ± 25.6	109.2±37.0	487.5±29.3	779.2±72.9
		88	114.2 ± 43.2	154.8±13.9	794.5±53.3	1289.3±162.9
2	Tensione limite sperimentale (MPa)	80	29.8 ± 5.1	21.1±4.2	131.8±9.2	126.4±16.1
		88	32.1± 5.5	23.1±4.7	142.7±14.9	113.5±7.8
3	Tensione limite media (MPa)		26.5 ± 4.9		128.5±12.2	
4	Forza di picco numerica (2D) $F_{num\ 2D}$ (N)	80	139.5	153.9	457.9	815.77
		88	137.7	183.6	653.1	1279.6
5	Errore relativo (%) $(F_{num\ 2D} - F_{exp})/F_{exp}$	80	39.8%	40.9%	-6.1%	9.9%
		88	17.0%	18.5%	-13.1%	3.1%
6	Forza di picco numerica (3D) $F_{num\ 3D}$ (N)	80			572.7	900.9
		88			682.5	1362.7
7	Errore relativo (%) $(F_{num\ 3D} - F_{exp})/F_{exp}$	80			17.4%	15.6
		88			-14.1%	5.6%

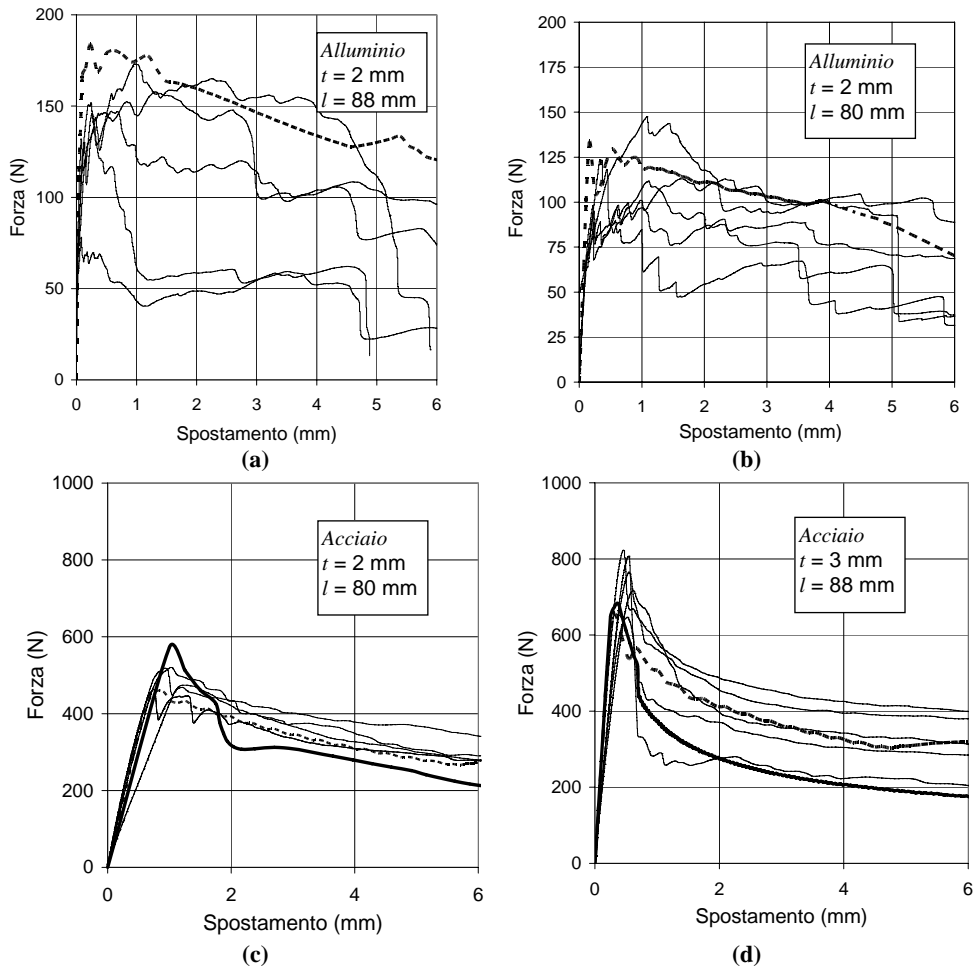


Fig. 3 – Curve sperimentali (linee sottili continue), numeriche 2D (linee spesse tratteggiate) e numeriche 3D (linee spesse continue) per due giunti di alluminio (a-b) e uno due di acciaio (c-d)

Dalla Figura 3 si osserva che i modelli piani agli elementi finiti predicono con grande accuratezza la risposta di tutti i giunti. In relazione ai valori di forza massima, l'errore relativo tra le risposte analitiche e numeriche raggiunge un massimo di circa il 41% per i giunti con aderendi di alluminio, mentre nel caso di giunti con aderendi di acciaio l'errore massimo è inferiore al 13% (Tabella 2).

Buona è anche l'accuratezza dei modelli tridimensionali agli elementi finiti. Relativamente ai valori di forza massima, l'errore relativo tra le risposte analitiche e numeriche raggiunge un massimo di circa il 17% (Tabella 2).

Sia i modelli computazionali piani che quelli tridimensionali, inoltre, colgono la risposta post-elastica con una buona precisione nel caso di giunti con aderendi di acciaio. La previsione è meno accurata, ma comunque concorde con il risultato sperimentale, per quelli di alluminio. La scadente previsione per i giunti con aderendi di alluminio è stata probabilmente causata dalla grande non linearità della risposta connessa a questi aderendi. Inoltre, si è osservato un cedimento decoesivo nei giunti di acciaio, mentre il cedimento si

è verificato all'interfaccia per certi giunti di alluminio. Questo può essere un'altra possibile causa della peggiore previsione dei giunti di alluminio. Per tali ragioni il modello computazionale 3D è stato sviluppato solo per le giunzioni con aderendi di acciaio.

Dai risultati precedenti il criterio di cedimento alle tensioni regolarizzate appare applicabile sia per calcolare con accuratezza la forza di picco che una giunzione può sostenere sia per esaminare il comportamento post-elastico del giunto fino al completo collasso. Anche l'approccio che sfrutta il criterio di cedimento della zona coesiva, seppur utilizzato solo per uno specifico materiale degli aderendi, dimostra una buona applicabilità, sia nella previsione elastica che post-elastica.

Il metodo computazionale proposto, che è stato precedentemente verificato come un metodo efficiente ed accurato per l'analisi delle tensioni in campo elastico [5], appare in grado di fornire buoni risultati anche per l'analisi della risposta post elastica di costruzioni incollate. Questa affermazione vale sia nel caso che si applichi un criterio di cedimento alle tensioni regolarizzate, che nel caso di un criterio di cedimento secondo il concetto di zona coesiva. L'efficienza del metodo in termini di riduzione del tempo di calcolo è stata di nuovo confermata, grazie alla drastica riduzione dei gradi di libertà del modello rispetto all'approccio classico [5-6].

In conclusione il modello computazionale proposto appare come una procedura semplice, efficiente ed accurata al fine di calcolare la risposta elastica e post-elastica fino a collasso di costruzioni incollate complesse.

5. CONCLUSIONI

Si è descritta una procedura efficiente basata su elementi finiti standard per l'analisi a collasso di giunzioni incollate. La procedura applica sia un criterio di cedimento alle tensioni regolarizzate (nel caso di modello piano) che un criterio di cedimento della zona coesiva (nel caso di modello 3D) ed è stata verificata, nel campo post-elastico, confrontandone le previsioni con le curve forza-spostamento di una campagna sperimentale basata su giunzioni T-peel.

L'accuratezza del modello, sia 2D che 3D, è apparsa molto buona in rapporto ai dati sperimentali, sia nel campo elastico, che nella previsione della forza di picco che nel comportamento post-elastico del giunto. Si è confermata inoltre la validità del criterio di cedimento alle tensioni regolarizzate ed anche l'efficienza della procedura in termini di tempo di CPU necessario per le analisi. La procedura dimostra viene proposta come un metodo semplice, efficiente ed accurato per l'analisi di costruzioni incollate molto complesse, tipiche dei settori automobilistico, ferroviario, navale e aeronautico.

Ringraziamenti

Si ringrazia il Dott. Piero Mauri di Henkel per il continuo e proficuo supporto all'attività sperimentale.

BIBLIOGRAFIA

- [1] Carlberger, T., and Stigh, U., 2007. "An Explicit FE-model of Impact fracture in an adhesive joint". *Engng. Fracture Mech*, **74**(14), pp. 2247-2262.
- [2] Goncalves, J. P. M., De Moura, M. F. S. F., et al, 2003. "Application of interface finite elements to three-dimensional progressive failure analysis of adhesive joints". *Fatigue Fract. Engng Master Struct*, **26**(5), pp. 479-486.

- [3] Hadavinia, H., Kawashita, L., Kinloch, A. J., Moore, D. R. and Williams, J. G., 2006. "A numerical analysis of the elastic-plastic peel test". *Engng. Fracture Mechanics*, **73**(16), Novembre, pp. 2324-2335.
- [4] Harris, J. A. and Adams, R. D., 1984. "Strength prediction of bonded single lap joints by non-linear finite element methods". *Int. J. Adhesion and Adhesives*, **4**(2), pp. 65-78.
- [5] Castagnetti, D., Dragoni, E., 2009. "Standard finite element techniques for the efficient stress analysis of adhesive joints". *Int J Adhes Adhes*, **29**, pp. 125-135.
- [6] Castagnetti, D., Spaggiari, A., Dragoni, E., 2007. "Metodi efficienti agli elementi finiti per l'analisi a collasso di strutture incollate". *Proceedings of the 36th AIAS*, Ischia (NA), ITALY, September 4-8.
- [7] Goglio, L., et al, 2008. "Design of adhesive joints based on peak elastic stresses", *Int J Adhes Adhes*, **28**, pp. 427-435.
- [8] Loctite – Hysol 9466, Technical Data Sheet, February 2006.
- [9] Bigwood, D. A., Crocombe, A. D., 1989. "Elastic analysis and engineering design formulae for bonded joints". *Int. J. of Adhesion and Adhesives*, **9**, pp. 229-242.
- [10] Pirondi, A, Moroni F, "Simulazione del cedimento di giunti ibridi rivet bonded attraverso modelli di danneggiamento" ABAQUS Regional Users' Meeting, Milano 4-6 Novebre 2008.
- [11] ABAQUS 6.7, 2006, "Users' Manual", HKS Inc.

PROGETTO A FATICA DI RUOTE DENTATE AD ELEVATE PRESTAZIONI

Alessandro Freddi

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia,

Università degli Studi di Bologna, Italia

E-mail: alessandro.freddi@unibo.it

Giorgio Olmi

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia,

Università degli Studi di Bologna, Italia

E-mail: giorgio.olmi@unibo.it

Abstract. *The present paper deals with the experimental analysis and the numerical modelling of the fatigue performance of gears. Such components, used for the mechanical transmission of a racing car, are made in high strength steel and carburized, quenched, shot-peened and superfinished before exercise. After a brief introduction on the state of the art on fatigue design methods, up to present main results are presented. The influence of shot-peening and of superfinishing on the fatigue limit were studied, then the operative parameters of the peening treatment were optimized, and the whole S-N curve of the peened component was determined. On the basis of DOE statistical methods and of theoretical models, a methodology for fatigue design of highly loaded components was developed.*

Keywords: *gears, bending fatigue, shot-peening, DOE, relative stress gradient*

1. INTRODUZIONE

La rottura delle ruote dentate si verifica spesso per effetto di fenomeni di fatica, che possono essere ascrivibili al progetto della ruota, alle condizioni di lavoro, alle caratteristiche del materiale. Il progetto a fatica è un tema sempre di grande attualità, per incrementare le capacità di carico del componente in questione e l'affidabilità generale del gruppo in cui è inserito. E' noto [1-2] che la maggior parte dei fenomeni di fatica di flessione e di contatto, si innescano in superficie o in prossimità di essa. Per questa ragione si è soliti adottare trattamenti superficiali, che, modificando le caratteristiche del materiale sulla strato più esterno, migliorano le prestazioni di fatica. Fra questi, si possono citare la rettifica e trattamenti di tipo termo-chimico, i quali hanno in comune la tendenza a rimuovere lo strato più esterno, che potrebbe essere sede di innesco. In [3] si è valutato anche l'influsso della nitrurazione gassosa, un trattamento che conferisce un'elevata durezza superficiale.

E' inoltre noto [4] che anche il livello delle tensioni residue ha indubbiamente un ruolo sulla resistenza a fatica dei componenti: è in quest'ottica che si inquadrano trattamenti come la pallinatura controllata o la sabbiatura, che conferiscono elevati valori di tensioni residue

di compressione sia in superficie, che immediatamente sotto di essa. Se è ormai riconosciuto il suo ruolo nel migliorare la resistenza a fatica, il risultato non è ancora facilmente correlabile ai parametri operativi effettivamente utilizzati. Diversi sono i fattori in gioco, fra cui il diametro dell'ugello di lancio dei pallini, la pressione, la portata, l'angolo di impatto, variabili inglobate nel parametro dell'intensità Almen. Tuttavia in [6] viene osservato che tale parametro non è di per sé sufficiente a spiegare il profilo di tensioni residue generate nel trattamento: per questa ragione è stato più volte indicato [2-3,5] nel diametro del pallino un ulteriore parametro da considerare per definire il processo. In [2] viene indagato il meccanismo di generazione delle tensioni residue, attribuibile a tre fenomeni. Di questi, i primi due [7] sono imputabili allo scorrimento plastico delle fibre superficiali e alle deformazioni plastiche al di sotto della superficie dovute alla pressione Hertziana. Il terzo fenomeno si manifesta sui componenti cementati e temprati e consiste nella trasformazione dell'austenite residua in martensite, con conseguente dilatazione volumica localizzata ed ulteriore incremento delle tensioni di compressione sottopelle.

Per quanto riguarda l'influenza dei parametri sul profilo tensionale residuo generato, in [2,5] si osserva che l'intensità Almen ha influenza essenzialmente sull'intensità del picco, mentre il diametro del pallino incide sulla sua collocazione in profondità, mostrando come pallini più grandi facciano sentire la propria azione più a fondo. Tali legami sono osservati empiricamente e non è tuttavia chiarito se vi siano relazioni lineari o non lineari fra i parametri operativi e le caratteristiche del profilo. Per quanto riguarda l'effetto sulla fatica flessionale, un aspetto spesso osservato è che la presenza della pallinatura porta l'innescò della cricca sottopelle. Tale aspetto, citato inizialmente da Wang [8] e poi confermato anche in [3,9-12] è da attribuirsi all'effetto delle tensioni residue. Mentre la compressione blindo la superficie, le tensioni residue di trazione sottopelle, rendono tali strati più vulnerabili, anche se meno sollecitati dai carichi esterni. E' inoltre spesso osservato che il ruolo della pallinatura e della distribuzione tensionale da essa generata sta essenzialmente nel ritardare la propagazione, piuttosto che nel prevenire l'innescò [5,10-13].

L'incremento sul limite di fatica che si può conseguire dipende dai parametri operativi, ma il suo valore massimo è legato al tipo di materiale trattato, ad esempio in [9] la sperimentazione riguarda una lega di Titanio e l'incremento non è molto considerevole, intorno al 6%, in [3,14] sono stati conseguiti incrementi fino circa al 20%, che rappresenta il valore più frequente. Nel caso di materiali con caratteristiche meccaniche inferiori, si possono conseguire incrementi maggiori, come in [1]: l'aumento massimo, intorno a 76% viene ottenuto, ottimizzando i parametri operativi. Le modalità secondo le quali viene condotta tale ottimizzazione, tuttavia, non sono in genere specificate; solo in [3] viene diffusamente utilizzato il metodo DOE (Design of Experiment), per pianificare le prove sperimentali, elaborare statisticamente i dati e ricercare l'optimum in maniera rigorosa.

Un'ulteriore opzione del trattamento consiste nell'operare una "ri-pallinatura" o una pallinatura aggiuntiva. Mentre la prima è di solito operata nel settore dell'aeronautica, per porre rimedio alla redistribuzione degli sforzi residui nei componenti, a seguito dell'esercizio e delle forti sollecitazioni termiche, oltre che meccaniche [9], la seconda è solitamente praticata [1,15] per aumentare l'intensità delle tensioni residue, soprattutto sullo strato superficiale. L'effetto della doppia pallinatura sul limite di fatica risulta controverso e non molto indagato, alcune prove riportate in [1] mostrano un effetto poco apprezzabile.

Un ulteriore trattamento superficiale, il cui effetto non è ancora stato ampiamente studiato è quello della tribofinitura (Isotropic Superfinishing, ISF). Tale trattamento è basato sull'azione congiunta di un reagente chimico e del moto vibratorio di sfregamento di

un mezzo non abrasivo, solitamente ceramico. A seguito delle operazioni, è possibile ottenere una superficie molto liscia, con ridottissima rugosità ($0,025 \div 0,050 \mu\text{m}$) e risvolti positivi nella riduzione di attrito ed usura. Risultati molto positivi nell'abbattimento od addirittura nell'eliminazione dei problemi di fatica per contatto e di pitting sono indicati in [16-18]. In [16] viene osservato che il trattamento ISF è particolarmente indicato, se praticato dopo quello di pallinatura: se da un lato è provato che la distribuzione di tensioni residue non è alterata dalla tribofinitura, dall'altro a seguito del pallinatura si generano delle microindentazioni superficiali, che, migliorando la lubrificazione, esaltano le proprietà di riduzione degli attriti prima enunciate. Esperienze condotte in [19] hanno inoltre sancito che tale riduzione di attrito ha un effetto benefico anche sulla fatica flessionale, limitatamente ad acciai a resistenza e durezza medio-basse. Nel caso degli acciai ad elevata resistenza, si segnalano solo pochi risultati riportati in [16], successivamente richiamati.

Alcuni modelli simulativi sono stati proposti per spiegare i risultati sperimentali o per prevedere le prestazioni a fatica di organi pallinati [2,5,10]. In [2] si è inoltre ipotizzato un legame lineare fra il picco di tensione residua ed il valore della resistenza a fatica. Secondo la teoria di Wang [8] le maggiori criticità sarebbero da ricercarsi in corrispondenza degli strati sub-superficiali, dove si hanno, nella distribuzione tensionale complessiva, le maggiori tensioni di trazione. Altri modelli specifici per i componenti pallinati sono stati proposti da Fuchs [20] e da Fatallah [11]. Se però il primo si presenta troppo rudimentale, essendo basato sull'uso del diagramma di Haigh e presupponendo la monoassialità degli sforzi, il secondo è analiticamente troppo pesante e di ardua applicazione anche per la difficile reperibilità dei dati sui materiali e soprattutto sulla difettosità dei componenti. Fra i modelli classici della fatica, si possono citare quelli che fanno uso delle curve S-N locali, differenziate per diverse gamme di coefficienti di concentrazione delle tensioni e di indici di effetto intaglio. Tuttavia tali modelli perdono di rigore, se applicati a componenti di forma non De Sanvenantiana, per i quali non sono correttamente definibili i coefficienti prima citati. Alcuni modelli sono stati proposti [21], per correlare la resistenza a fatica al gradiente di tensione. Uno dei più recenti, sviluppato da Eichlseder [22], è basato sull'applicazione combinata di studi agli elementi finiti per la stima del gradiente e di equazioni di tipo empirico. Un altro approccio sulla fatica è quello proposto da Taylor nella sua recentissima teoria delle distanze critiche [23].

Il presente lavoro riguarda una ricerca sulla progettazione a fatica flessionale di ruote dentate. Queste ultime sono impiegate nella trasmissione meccanica di auto da competizione e sono realizzate in acciaio ad elevata resistenza con caratteristiche prestazionali molto spinte. Tale lavoro affronta, anche con metodi statistici, diversi aspetti, fin ad ora non pienamente indagati, quali l'influsso dei parametri operativi sulla distribuzione di tensioni residue, l'influenza della doppia pallinatura e della tribofinitura sulla resistenza a fatica, un'indagine ad ampio spettro per l'ottimizzazione dei trattamenti di pallinatura, individuando il parametro operativo più significativo. Il presente lavoro si propone inoltre di discutere i risultati in chiave di metodologie di progettazione a fatica, anche con l'uso di modelli fino ad ora non comunemente applicati al caso di corpi pre-pallinati.

2. MATERIALI E METODI

Il componente studiato, Fig. 1, è una ruota dentata a denti dritti: questa è realizzata in modo standard, conformemente con le specifiche industriali dell'azienda, ma è leggermente

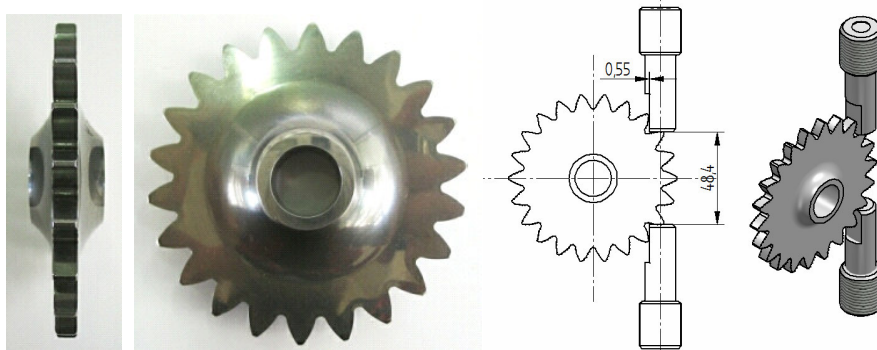


Figura 1. Componente sottoposto alle prove di fatica e riferimento ai punzoni

rimpicciolita rispetto al componente tradizionale. Il materiale è un acciaio al Ni-Cr ad elevata resistenza, con trattamenti termici e superficiali, per incrementarne ulteriormente la resistenza a fatica. Si attuano nell'ordine la cementazione, la tempra, la rettifica, la pallinatura e la tribofinitura. In [24] viene rimarcato il fatto che la pallinatura dà buoni risultati sugli acciai ad elevata resistenza, dal momento che la distribuzione di tensione residua si mantiene solitamente stabile in esercizio. Inoltre sono molteplici le applicazioni su acciai cementati [1,25-26]: in tal caso la pallinatura dopo tempra ha un effetto amplificato dalla conversione dell'austenite residua.

Per le prove di fatica descritte, oltre che per i rilievi diffrattometrici delle tensioni residue dopo successivi attacchi chimici con dissoluzione elettrolitica di strati successivi di materiale, si sono utilizzate in totale 30 ruote con le seguenti caratteristiche: 22 denti, modulo pari a 4 mm, larghezza di fascia di 5,5mm, fattore di correzione, 0,96, addendum pari a 3,73 mm e dedendum uguale a 5,29 mm.

Le prove sperimentali sono state realizzate su un vibroforo risonante Rumul con l'impiego di un'ideale attrezzatura [27-28], di cui si riporta un'immagine in Fig. 1. Tale attrezzatura permette l'esecuzione di prove di "tooth bending fatigue", ponendo sotto carico una coppia di denti sollecitati da due punzoni gemelli, posti l'uno superiormente e l'altro inferiormente. Cinque coppie utili per ciascuna ruota sono state individuate, coerentemente con i requisiti di montaggio sull'attrezzatura e avendo cura di non sollecitare i denti adiacenti. I punzoni sono realizzati in acciaio per utensili con elevata durezza e resistenza all'usura, rispettando tolleranze geometriche e dimensionali strette. Queste consentono ad esempio di garantire il parallelismo fra le facce che trasmettono il carico flessionale ai denti e quelle accoppiate al bancale e alla traversa della macchina, assicurando così il corretto allineamento. L'attrezzatura ed il relativo ciclo di montaggio sono progettati in modo tale da sottoporre a carico solo i due denti interessati dalla prova, lasciando scarichi tutti gli altri ed in particolare quelli adiacenti. I risultati delle prove di fatica sono stati elaborati con il metodo Dixon Staircase abbreviato [29], già applicato anche in [3,5,10].

Le tensioni residue dovute alla pallinatura sono state misurate con la tecnica della diffrazione a raggi X [2-3,5,24,30] ed i loro valori sono stati stimati sulla base del metodo "sin²ψ" con radiazione CrKα sul piano {211} della fase bcc (XStress3000, Stresstech, Finland). La valutazione delle tensioni residue è stata eseguita al piede del dente, rimuovendo gli strati più superficiali del materiale. I valori così determinati sono stati corretti tramite apposito algoritmo [31], per tenere conto dell'avvenuta rimozione di

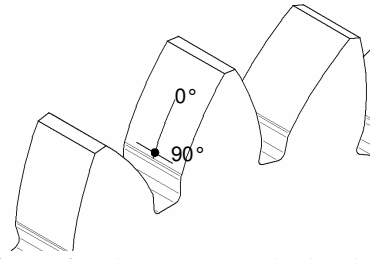


Figura 2. Misura delle tensioni residue

materiale. In corrispondenza della radice del dente le tensioni residue sono state determinate lungo due direzioni, 0° e 90° , rispettivamente perpendicolare e parallela all'asse della ruota (Fig. 2). I risultati hanno mostrato come i valori lungo tali direzioni sono molto prossimi fra loro, il che ha autorizzato a considerare una distribuzione media nel prosieguo del lavoro. Si è cioè considerata una distribuzione residua equibiassiale, anche in accordo con [32]. L'analisi di tensioni residue è stata effettuata in alcuni casi prima dell'esecuzione delle prove di fatica (alcune ruote sono state sacrificate solo ai fini della presente valutazione), in altri su ruote già sottoposte a prova, avendo cura di effettuare i rilievi su denti non sollecitati. Tale protocollo è stato seguito, per tutelarsi dai fenomeni di rilassamento delle tensioni residue a seguito dell'esercizio descritti in [11,24,33-35].

3. RISULTATI SPERIMENTALI E DISCUSSIONE

La campagna sperimentale è stata articolata in diverse fasi, ognuna delle quali è stata supportata da analisi di tensioni residue, da valutazioni statistiche con strumenti DOE e da indagini a carattere numerico-simulativo. Le fasi possono essere così riassunte:

- Valutazione dell'influsso dei trattamenti di pallinatura e tribofinitura sul comportamento a fatica del componente in esame.
- Determinazione della curva S-N locale (tratto inclinato e tratto orizzontale) in relazione al componente pallinato e tribofinito.
- Analisi comparata delle prestazioni a fatica a seguito di diversi trattamenti di pallinatura, organizzati secondo un piano a due fattori (intensità Almen e diametro del pallino).

Influsso dei trattamenti di pallinatura e di tribofinitura

Il primo problema che è stato posto ha riguardato la valutazione in termini quantitativi dell'influsso del trattamento di pallinatura sulla vita a fatica del componente in esame. Si sono predisposte due ruote dentate cementate, temprate, rettificate e tribofinite e altre due che avessero subito il trattamento di pallinatura prima della tribofinitura. Tale trattamento è stato operato con pallini S110 (del diametro di circa 0,3 mm) lanciati alla pressione di 4 bar.

In Fig. 3 si mostrano i risultati dei rilievi delle tensioni residue: si nota l'influsso della pallinatura nella generazione di un picco di tensione, che si va a sovrapporre alla distribuzione residua dovuta essenzialmente ai pregressi trattamenti termici.

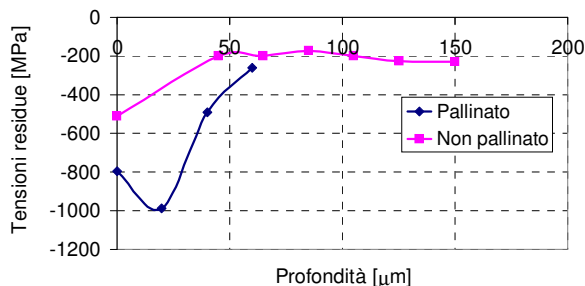


Figura 3. Analisi di tensioni residue: pallinato vs. non pallinato

I risultati delle prove a fatica sono riassunti nelle Tabb. 1 e 2 [27-28,36]: ogni sessione di prova è stata condotta fino a rottura (segno X in tabella) o fino al raggiungimento di 10^7 cicli (evento di non rottura, segno O). I valori di carico nella prima colonna si riferiscono alla massima forza F_{max} esplicitata dalla coppia di punzoni sui denti coinvolti nella prova. Per il carico si è considerata una legge pulsante, con l'aggiunta di un leggero precarico (fissato a 1,25 kN), rivelatosi necessario, per garantire il mantenimento del corretto riferimento della ruota all'attrezzatura.

Tabella 1. Risultati delle prove sulla ruota non pallinata (tribofinita)

F_{max} [kN]	1	2	3	4	5	6
9.50	X		X			
9.25		O		X		
9.00					X	
8.75						O

Tabella 2. Risultati delle prove sulla ruota pallinata (e tribofinita)

F_{max} [kN]	1	2	3	4	5	6
11.00	X		X			
10.75		O		X		O
10.50					O	

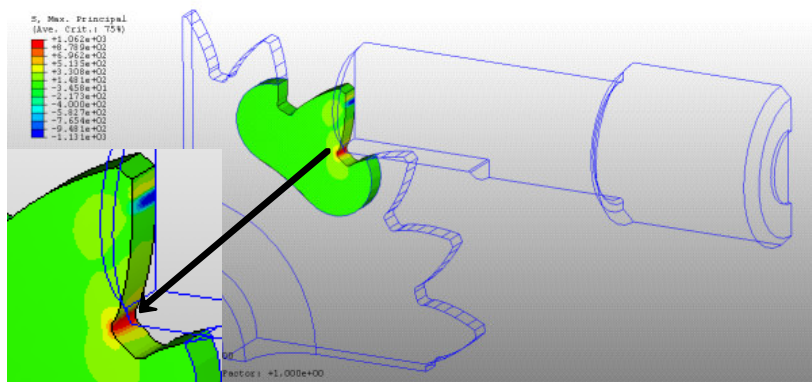


Figura 4. Modello FEM per relazionare il carico applicato al picco di tensione al piede

Tabella 3. Risultati delle prove sulla ruota pallinata e non tribofinita

F _{max} [kN]	1	2	3	4	5	6	7	8
12.00							X	
11.75						O		X
11.50					O			
11.25				O				
11.00	X		O					
10.75		O						

I carichi limite a fatica sono stati calcolati in accordo con le formulazioni del metodo Dixon [29] per un numero di prove nominali pari a sei. Per conferire generalità ai risultati ottenuti, essi sono stati convertiti in limiti a fatica locali in termini di tensione al piede del dente. Per effettuare tale conversione, ci si è avvalsi di un modello statico agli elementi finiti [28], opportunamente realizzato per stimare il picco di tensione flessionale in funzione del carico trasmesso dai punzoni. Ai fini della simulazione si è considerato un modello formato da un quarto della ruota, discretizzato con l'uso di elementi solidi ad otto nodi ad integrazione ridotta, attivando il controllo sui modi deformativi a clessidra. La trasmissione del carico è stata impostata in corrispondenza del contatto fra il quarto di ruota ed un punzone rigido, anch'esso opportunamente modellato (Fig. 4).

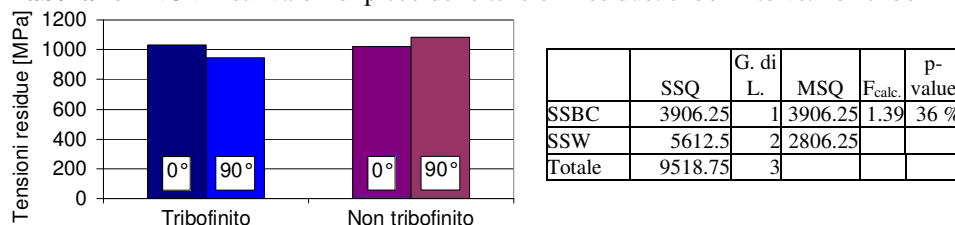
Il modello è completato da un sistema di vincolamento che lascia libere solo le rotazioni attorno all'asse del quarto di ruota. Si è così ricavato che a un carico di 10 kN corrisponde una tensione di picco di 1115 MPa: interpolando sulla base di questo valore si sono ricavati i limiti di fatica locali per ogni casistica affrontata nel presente lavoro.

Per il componente non pallinato si è ottenuto un limite di 1010 MPa, mentre nel caso della pallinatura la resistenza è risultata incrementata del 19%, arrivando a 1203 MPa. L'incremento prestazionale è perfettamente in linea con altri risultati ottenuti in precedenti ricerche [3,14].

Seguendo la medesima impostazione, si è valutato l'influsso della tribofinitura praticata dopo la pallinatura. Si è quindi considerata un'ulteriore coppia di ruote dentate, sottoposte ai trattamenti di cementazione, tempra, rettifica e pallinatura (secondo i parametri operativi indicati al punto precedente), ma non tribofinita. Queste ruote sono state provate a fatica secondo la medesima procedura ed i risultati sono stati posti a confronto con quelli già ottenuti per le ruote pallinate e tribofinite. Anche nel presente caso si sono effettuati rilievi di tensioni residue, per valutare se l'esecuzione della tribofinitura potesse intaccare il profilo tensionale dovuto alla pallinatura. Si riportano in Tab. 3 i risultati delle prove di fatica.

Per quanto riguarda l'eventualità che la tribofinitura possa avere un effetto sull'alterazione del profilo di tensioni residue dovuto alla pallinatura, si può citare il lavoro di Winkelmann [16], che esclude questa possibilità e di conseguenza anche quella di un influsso sfavorevole sulla resistenza a fatica. Nel presente caso si osserva (Tab. 4) come il livello di tensioni residue sulla ruota pallinata e tribofinita sia leggermente inferiore rispetto a quanto rilevato sulla ruota non tribofinita. Tuttavia le differenze sono molto limitate: i valori di picco (che secondo il modello in [2] sono proporzionalmente legati alla resistenza a fatica) sono molto vicini. Oltre a questo va anche rilevato il fatto che i dati provengono da due ruote diverse, sottoposte ai medesimi trattamenti termici e di pallinatura, ma in momenti diversi. Quindi all'incertezza insita nel procedimento di misura si aggiungono quelle dovute alla varianza dei trattamenti. Tali considerazioni sono state confermate da

Tabella 4. ANOVA sui valori di picco delle tensioni residue: tribofinito vs. non tribofinito



	SSQ	G. di L.	MSQ	F _{calc.}	p-value
SSBC	3906.25	1	3906.25	1.39	36 %
SSW	5612.5	2	2806.25		
Totale	9518.75	3			

Tabella 5. Analisi ANOVA sui valori delle resistenze a fatica: tribofinito vs. non tribofinito

	SSQ	v _i	MSQ	F _{calc.}	p-value
Fattore tribofinitura	2924	1	2924	1.82	20%
Errore	---	12	Stimato: s _p ² = 1608		
Totale	2924	13			

un'analisi ANOVA (ANalysis Of VARIance) [37], che confronta i valori di picco e considera come due repliche i valori rilevati lungo le direzioni 0° e 90° illustrate in Fig. 2. I risultati sono riassunti nella Tab. 4.

Venendo al limite di fatica locale, quello della ruota non tribofinita ammonta a 1280 MPa, da porre a confronto con quello di 1203 MPa per la ruota tribofinita precedentemente calcolato. Si può osservare che, a seguito della tribofinitura il limite di fatica è leggermente diminuito (di circa il 6%). In accordo con [29] sono stati calcolati errori e deviazioni standard: queste ultime sono risultate pari a s₁=21,0 MPa nel caso del componente tribofinito ed a s₂=49,4 MPa in quello del componente solo pallinato. Come suggeriscono anche le successioni degli eventi di rottura e non rottura riportate nelle Tab. 2-3, l'aggiunta della tribofinitura riduce fortemente la dispersione dei risultati, diminuzione che si traduce in un sensibile abbattimento della deviazione standard (-57%). L'analisi ANOVA (Tab. 5) è stata condotta secondo [37], stimando l'errore come suggerito in [3]. Dal momento che le prove effettuate sono in totale 14 (6 per la stima del limite del componente pallinato e tribofinito e 8 per la determinazione della resistenza nell'altro caso), un totale di 13 (14-1) gradi di libertà è disponibile per l'analisi. Di questi v₁ + v₂ = (6 - 1) + (8 - 1) = 12 sono assegnati all'incertezza, mentre 1 è attribuito all'effetto del "fattore tribofinitura". Sulla base di questo, l'errore può essere stimato come media pesata delle varianze sulla base dei gradi di libertà (Eqn. (1)).

$$s_p^2 = \frac{s_1^2 \cdot v_1 + s_2^2 \cdot v_2}{v_1 + v_2} = \frac{21.0^2 \cdot 5 + 49.4^2 \cdot 7}{5 + 7} = 1608 MPa^2 \quad (1)$$

Il test statistico (Tab. 5) ha confermato che non si osservano differenze significative alla soglia del 5%, cioè che i due limiti trovati sono statisticamente gli stessi. E' quindi da escludere un influsso della tribofinitura sul limite di fatica, fermo restando il ruolo benefico nel migliorare la ripetibilità dei risultati. Tali osservazioni trovano piena conferma da

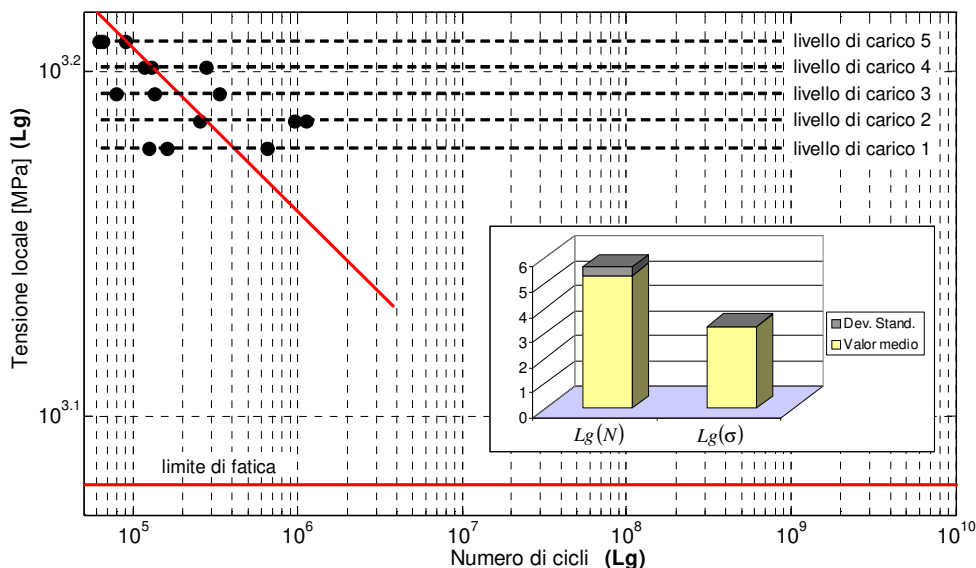


Figura 5. Curva S-N ricavata sperimentalmente ed analisi sulla dispersione dei risultati

ricerche simili descritte in [16], sia per quanto concerne la trascurabile influenza sul limite di fatica, sia per quanto riguarda la drastica riduzione (diminuita fino ad 1/3) della deviazione standard.

Determinazione della curva S-N locale per il componente pallinato e tribofinito

Attività sperimentale. In accordo con ISO 12107 [38] si è determinata la curva locale P-N (carico-numero di cicli) in un intervallo per cicli da 10^5 a 10^6 , una tipica durata per le trasmissioni delle auto da competizione. In base a prove preliminari si era riscontrato che un tale intervallo di vita finita corrispondeva ad un carico massimo tra 13,50 e 14,50 kN. Secondo [38] tale range è stato suddiviso in cinque livelli equispaziati: 13,50 kN; 13,75 kN; 14,00 kN; 14,25 kN; 14,50 kN. Le prove sono state condotte con tre ripetizioni ad ogni livello di carico per un totale di 15 prove. In tutte queste prove si sono riscontrati solo eventi di rottura, determinando la pendenza della curva in accordo con [38]. I risultati sperimentali, relativi agli eventi di rottura in funzione del carico applicato rispetto al numero di cicli, sono stati poi rielaborati, pervenendo alla curva S-N locale (Fig. 5).

Il modello agli elementi finiti descritto al paragrafo precedente è stato a tal fine utilizzato, per operare la conversione fra carico applicato e tensione di picco. Il digramma determinato considera la tensione principale massima (le tensioni principali sono molto vicine l'una all'altra, per cui, in stato piano di tensione, quasi coincidono con la tensione ideale di Von Mises). L'equazione della retta interpolante è stata determinata, verificando preventivamente l'adeguatezza del modello lineare. La dispersione dei risultati è stata infine discussa, calcolando le deviazioni standard delle distribuzioni logaritmiche di tensione ($\text{Log } \sigma$) e di cicli ($\text{Log } N$). Nella Fig. 5 l'istogramma mostra la proporzione fra le deviazioni standard ed i valori medi: i rapporti fra le due grandezze sono molto bassi (circa il 6% per

Log N, meno dell'1% per Log σ). La curva determinata si riferisce ad una probabilità di rottura del 50%, ed il numero delle prove e delle ripetizioni è stato scelto sulla base di [38], per avere una confidenza prossima al 90%.

Possibilità di simulazione. Sono state messe a punto delle tecniche che permettessero di simulare il trattamento di pallinatura, con particolare riferimento al fenomeno di impatto ed alle tensioni residue di compressione generate. In [7] si è del resto messo in evidenza come il fenomeno di urto dinamico influenzi in maniera significativa la distribuzione tensionale indotta. Nella simulazione degli impatti, risultati apprezzabili si possono avere con l'applicazione dell'algoritmo di integrazione esplicita. Si è infatti osservato [39] che tale tecnica non richiede il calcolo della matrice di rigidezza tangente, così che ogni passo di calcolo è reso decisamente più rapido rispetto al caso di integrazione implicita. Un'analisi di sensitività è stata orientata a valutare quanto il risultato finale sia influenzato dalla rigidezza del pallino. E' emerso come risultati accettabili si possano avere, supponendo il pallino infinitamente rigido, il che è motivato dalla grande differenza fra la rigidezza di questo e quella dell'organo bersaglio. La simulazione del contatto è stata realizzata con un algoritmo "weighted master-slave", imponendo come condizione iniziale la velocità dei pallini in fase di lancio verso la superficie. I risultati ottenuti sono mostrati in Fig. 6 con riferimento ai due modelli realizzati in due e tre dimensioni. Le differenze esistenti fra la curva sperimentale e quelle simulate possono essere imputate alle tensioni residue generate durante la tempra, che non possono essere considerate dal modello di simulazione [40-41].

Come anticipato, il modello del gradiente lega il limite di fatica al gradiente di tensione relativo, così definito (Eqn. (2)):

$$\chi' = \left(\frac{1}{\sigma_{max}} \right) \left(\frac{d\sigma}{dx} \right) \quad (2)$$

Il limite a fatica locale σ_f in corrispondenza di un generico punto nel volume di un componente, caratterizzato da un valore di tensione σ_{max} , può essere predetto con la stima del gradiente di tensione relativo nel punto. Per l'applicazione in studio si è utilizzata l'Eqn. (3) [22], che permette di ricavare il limite locale, interpolando rispetto a dati sui limiti a fatica di provini lisci del medesimo materiale, soggetti a trazione - compressione (σ_{tf}) e a flessione (σ_{bf}).

$$\sigma_f = \sigma_{tf} \cdot \left(1 + \left(\frac{\sigma_{bf}}{\sigma_{tf}} - 1 \right) \left(\frac{\chi'}{(2/b)} \right)^{K_D} \right) = \sigma_{tf} \cdot n_x \quad (3)$$

La classica equazione che descrive il tratto inclinato della curva S-N locale è la seguente (Eqn. (4)):

$$N = N_D \left(\frac{\sigma_f}{\sigma} \right)^k \quad (4)$$

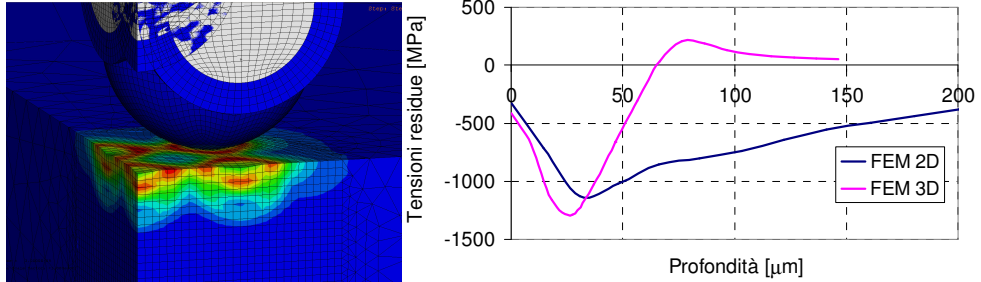


Figura 6. Simulazione degli impatti dei pallini e della distribuzione residua indotta

La pendenza k e il numero di cicli al ginocchio N_D sono stati calcolati secondo le seguenti formule (Eqn. (5) ed Eqn. (6)), che completano il modello di Eichlseeder: si osserva che anche tali grandezze dipendono dal gradiente relativo, (attraverso il coefficiente n_x).

$$\text{Log}(N_D) = \text{Log}(N_{Dmin}) + \frac{\text{Log}(N_{Dmax}) - \text{Log}(N_{Dmin})}{n_x^{k_n}} \quad (5)$$

$$k = k_{min} + \frac{k_{max} - k_{min}}{n_x^{k_k}} \quad (6)$$

K_D , k_n , and k_k sono coefficienti caratteristici per il materiale del componente riportati in [22]. Sulla base delle Eqn. (4), (5) e (6), per definire σ_f , k ed N_D , è richiesta la conoscenza di due set di dati sperimentali, sostanzialmente due curve S-N per provini privi di intaglio in due distinte condizioni di carico. Generalmente si richiedono una curva S-N per prove condotte con gradiente relativo nullo (da cui si determinano i parametri N_{Dmax} , numero di cicli al ginocchio e k_{max} , pendenza del tratto inclinato) ed una per prove con gradiente χ' elevato (si ricavano N_{Dmin} e k_{min}). Solitamente si fa riferimento, come anticipato nell'equazione (3) a dati a trazione-compressione ed a flessione, per i quali si ha $\chi' = 2/b$, essendo b lo spessore o il diametro del provino.

La distribuzione tensionale totale (al carico massimo) è stata ricavata, sommando il contributo delle tensioni residue, ricavato numericamente, a quello degli sforzi da carichi esterni, seguendo l'impostazione riportata in [8]. Si è parallelamente tenuto conto del fatto che la distribuzione tensionale da carico esterno varia ciclicamente, mentre quella dovuta alla pallinatura rappresenta un termine medio che si può supporre costante, a meno di fenomeni di redistribuzione. La Fig. 7 mostra inoltre il valore di χ' alla radice del dente, operando un raffronto in presenza ed in assenza di pallinatura: si può osservare come il gradiente di tensione relativo (38 mm^{-1}) sia fortemente incrementato (di un fattore intorno a 5) grazie al trattamento di pallinatura. Il metodo del gradiente fornisce quindi un'ulteriore interpretazione dell'effetto benefico della pallinatura: tale trattamento, oltre a generare tensioni residue di compressione, che abbassano lo stato tensionale totale, comporta un forte incremento del gradiente relativo di tensione nello strato superficiale, che permette di prevenire in parte anche l'innesco. Il limite a fatica locale (all'inversione) è stato quindi valutato sulla base dell'Eqn. (3). Si è tenuto conto delle tensioni residue come tensioni medie, come suggerito in [42], supponendole invariate nel corso della vita del componente

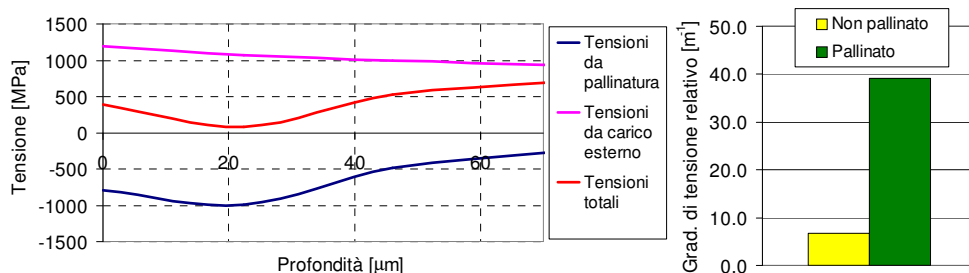


Figura 7. Determinazione della distribuzione tensionale totale ed influsso della pallinatura sull'incremento del gradiente di tensione relativo

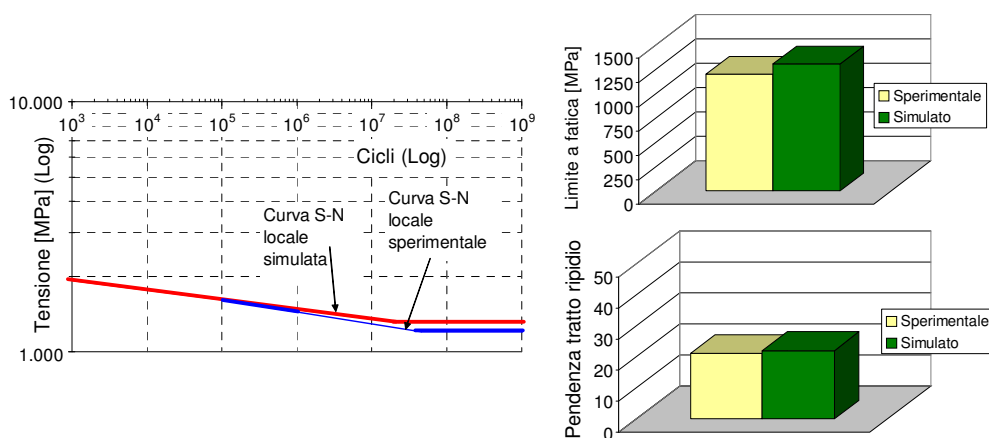


Figura 8. Determinazione della curva S-N e confronto fra risultati numerici e sperimentali

[32]. La resistenza all'escursione ed il limite locale sono stati successivamente determinati in base alla tensione media, tramite costruzione del diagramma di Haigh [43-44]. Si è quindi tenuto conto delle tensioni residue da pallinatura in due sedi: per calcolare la distribuzione di tensione effettiva, sulla base della quale stimare il gradiente relativo, e per operare l'aggiustamento del limite anzidetto.

Attraverso le Eqn. (5-6) si è infine pervenuto alla totale simulazione della curva S-N locale. Nella Fig. 8 la curva è rappresentata unitamente a quella sperimentale, evidenziando un buon accordo. L'istogramma mette a confronto il limite di fatica e la pendenza: gli errori riscontrati ammontano al 4% (sulla pendenza) ed al 9% (sul limite di fatica).

Analisi comparata delle distribuzioni di tensione residua e dei limiti a fatica per diversi trattamenti di pallinatura

Una volta apprezzati gli effetti di pallinatura e tribofinitura e ampiamente studiato il comportamento a fatica a seguito del trattamento comunemente praticato, si è voluto valutare come tale trattamento potesse essere ottimizzato. Si è cioè voluto indagare come l'adozione di diversi parametri operativi nella pallinatura del componente in esame potesse

Tabella 6. Piano a due fattori dei trattamenti indagati

Diametro pallino Intensità Almen	S110 (0,28 mm)	S170 (0,43 mm)	S230 (0,58 mm)	S330 (0,84 mm)
Bassa (press. \cong 4 bar)	●	●	●	●
Media (press. \cong 4,5 bar)	●	●	● ♦	♦
Alta (press. \cong 5 bar)	-	-	♦	♦

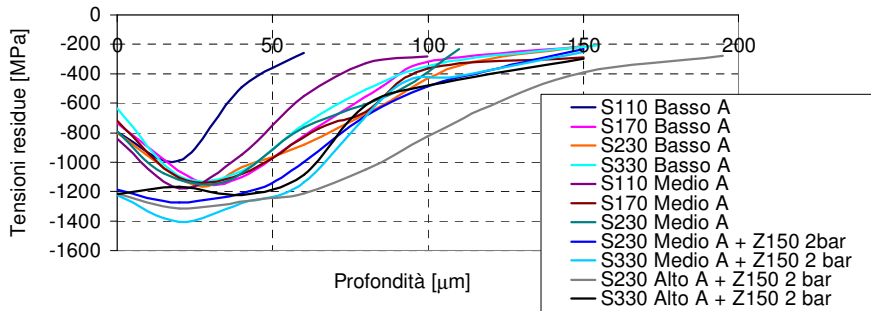


Figura 9. Analisi di tensioni residue su 11 diversi trattamenti

influenzare la distribuzione tensionale residua indotta ed il limite di fatica. Il problema è stato anche in questo caso affrontato, partendo dall'analisi diffrattometrica, proseguendo con le prove sperimentali ed infine con le indagini numeriche.

Ai fini della presente analisi ulteriori 10 trattamenti (oltre al già considerato S110 Basso A con pressione di lancio di 4 bar) sono stati considerati (tutti con tribofinitura): essi sono stati organizzati in un piano a due fattori (diametro del pallino e intensità Almen), prevedendo anche trattamenti di doppia pallinatura. I parametri operativi sono riassunti nella Tab. 6, in cui i livelli dell'intensità Almen sono equispaziati, con pressione variante da 4 a 5 bar circa, mentre il diametro del pallino è indicato secondo la notazione convenzionale in decimi di pollice ed in mm. I trattamenti indicati con il simbolo (●) sono stati praticati, incrociando i parametri presenti nella matrice, quelli individuati dal segno (♦), oltre che essere eseguiti come indicato, hanno previsto anche un trattamento aggiuntivo con pallini ceramici, con un diametro di 0,15 mm, sparati ad una pressione di 2 bar. Tale trattamento addizionale (Z150, 2 bar) completa la doppia pallinatura.

Analisi diffrattometriche ed elaborazioni statistiche dei dati. La prima analisi effettuata ha riguardato le tensioni residue: le rispettive distribuzioni sono confrontate in Fig. 9.

Si può osservare come i profili di tensione residua abbiano, nel caso di pallinatura singola, il tipico andamento già osservato in molteplici occasioni, ad esempio in [1-2,10], con il valore di picco collocato sottopelle, ad una distanza dalla superficie abitualmente collocata fra 20 e 30 µm. La doppia pallinatura viene spesso attuata nel settore delle ruote dentate ad elevate prestazioni, essenzialmente per uniformare e ridurre la rugosità superficiale. Il trattamento aggiuntivo prima descritto si rende essenziale nel caso dei trattamenti più spinti, con elevato valore dell'intensità Almen e del diametro del pallino. Nel caso del trattamento S230 a media intensità Almen si sono valutati i due casi: con o

Tabella 7. Analisi della varianza sull'intensità del picco massimo

Picco massimo [MPa]		Diametro pallino			
		S 110	S 170	S 230	S 330
Int. Almen	Basso	-1034	-1093	-1214	-1111
		-946	-1157	-1099	-1080
	Medio	-1158	-1028	-1250	-1465
		-1201	-1184	-1300	-1350
	Alto			-1328	-1184
				-1300	-1260

	SSQ	G. di L.	MSQ	F _{calc.}	p-v. [%]
Int. Almen					
Almen	123024	2	61512	16.51	0.07
Diam. pallino	102893	3	34298	9.21	0.32
Interaz.	38664	4	9666	2.60	10.09
Errore	37248	10	3725		
Totale	301830	19			

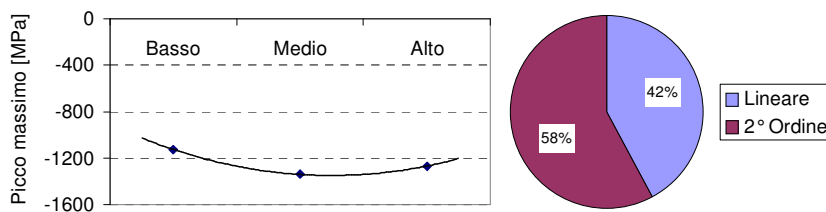


Figure 10. Analisi di linearità sulla dipendenza picco massimo – intensità Almen

senza il trattamento aggiuntivo, per apprezzare le differenze nella distribuzione residua e, successivamente, nella resistenza a fatica. Si nota come la seconda pallinatura porti ad un sensibile incremento della tensione residua superficiale e renda la curva sostanzialmente stabile nei primi strati sub-superficiali, rendendo il picco sottopelle molto meno evidente. Si osserva parimenti come poco apprezzabile sia l'effetto sul massimo di tensione residua, coerentemente con quanto rimarcato anche in [1]. Anche gli altri trattamenti di doppia pallinatura portano a profili qualitativamente simili, con una elevata tensione superficiale ed un primo tratto della curva quasi costante.

Per valutare la significatività dei fattori in esame nella definizione del massimo picco e della sua posizione, si è usato lo strumento dell'ANOVA [37]. L'analisi è stata effettuata, interpretando i due rilievi lungo le direzioni 0° e 90° (Fig. 2) come due ripetizioni e considerando per il caso S230 medio A il trattamento di doppia pallinatura, ossia quello che permetteva di soddisfare le specifiche sulla rugosità. L'analisi statistica (Tab. 7) mostra che entrambi i fattori, intensità Almen e diametro del pallino, hanno un impatto sul valore del picco (alla soglia di significatività dell'1%): di questi il fattore più significativo (valore più basso del p-value) è l'intensità Almen. Tale risultato è in accordo con quanto osservato in [2], si nota inoltre che la decomposizione della varianza effettuata esclude la presenza di interazione tra i due fattori. Considerando i valori delle ultime due colonne del piano fattoriale, si è inoltre affinata l'analisi, per indagare la natura del legame rispetto all'intensità Almen. Si è utilizzata allo scopo la tecnica della decomposizione ortonormale della somma quadratica [37], evidenziando la parte lineare e quella non lineare. La Fig. 10 mostra la ripartizione fra la componente lineare e quella del secondo ordine: ambedue risultano significative, ma è proprio la componente non lineare a prevalere. Tale dato è confermato dall'andamento del valore di picco rispetto all'intensità Almen (mediato sui due valori di diametro dei pallini), mostrato sulla sinistra. Il diagramma mostra come, all'aumentare dell'intensità Almen, il picco aumenti, con tendenza ad una saturazione.

Tabella 8. Analisi della varianza sulla profondità di picco

Profondità di picco [µm]		Diametro pallino			
		S 110	S 170	S 230	S 330
Int. Almen	Basso	20	30	26	30
		20	30	28	25
	Medio	22	35	20	20
		20	23	20	20
	Alto	X		20	40
				25	40

	SSQ	G. di L.	MSQ	F _{calc.}	p-v. [%]
Int. Almen	207	2	103	10.23	0.38
Diam. pallino	277	3	92	9.13	0.33
Interaz.	254	4	64	6.29	0.85
Errore	101	10	10		
Totale	838	19			

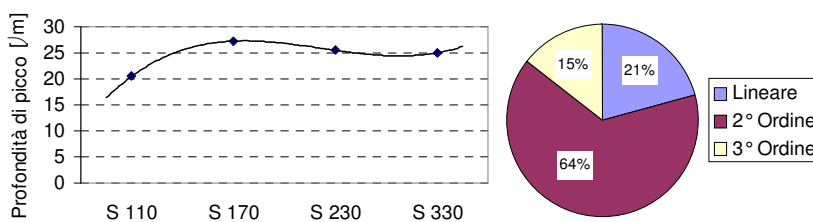


Figura 11. Analisi di linearità sulla dipendenza profondità di picco – diametro del pallino

Con la medesima impostazione è stata effettuata l'analisi della varianza in relazione alla profondità del picco. L'indagine statistica ha portato a concludere (Tab. 8) che entrambi i fattori hanno influenza sulla profondità del picco, e pure l'interazione è significativa alla soglia dell'1%. Il fattore più significativo è il diametro del pallino, coerentemente con quanto osservato in [2].

Considerando le prime due righe del piano fattoriale, ed applicando la medesima tecnica prima citata, si sono separate le componenti del primo, secondo e terzo ordine nel legame fra profondità di picco e diametro del pallino (Fig. 11). La componente prevalente e l'unica che risulta significativa è quella del secondo ordine. Il grafico (mediato sui valori dell'intensità Almen) mostra la forte non linearità con tendenza alla saturazione per valori elevati del diametro del pallino, anche la componente del terzo ordine, pur non significativa, ha comunque un certo effetto. Tali risultati dimostrano i limiti strutturali del trattamento di pallinatura: quanto più esso è reso spinto, tanto più comporta costi maggiori e una rugosità più elevata, senza che tali svantaggi siano compensati da un maggiore tensione residua o da una maggiore profondità di azione.

Risultati sperimentali delle prove di fatica ed elaborazioni statistiche. Per l'effettuazione delle prove di fatica sono state scartate alcune combinazioni di pallinature. Il riferimento è ai trattamenti S230 Basso A ed a tutti quelli con il pallino di dimensione maggiore (S330). Oltre che sulla base del profilo di tensione residua, il primo trattamento è stato scartato per via della scarsa diffusione, coniugando esso un pallino abbastanza grande con una bassa intensità Almen, tutti gli altri sono stati scartati perché la rugosità che conferivano alla superficie era inaccettabile, anche a seguito del trattamento aggiuntivo. Le prove, hanno fornito i seguenti risultati in termini di limite locale al piede del dente (Fig. 12) [27]. Nell'istogramma sono stati aggiunti anche il limite per la ruota non pallinata e

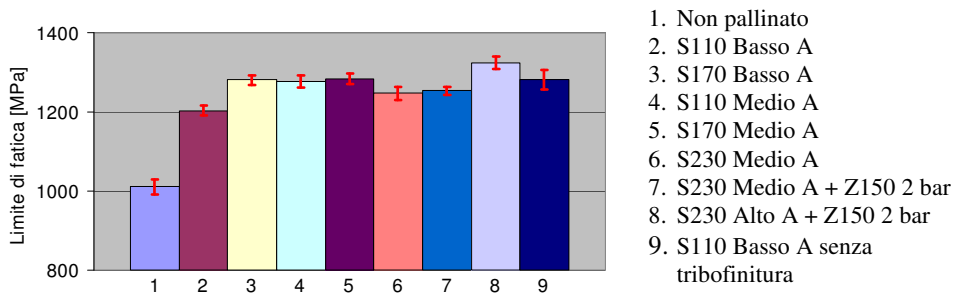


Figura 12. Limiti a fatica locali per i trattamenti studiati e relativa legenda

Tabella 9. Risultati delle prove sulla ruota pallinata S230 Medio A

F_{max} [kN]	1	2	3	4	5	6	7	8	9	10
11.75							X			
11.50						O		X		
11.25	X				O				X	
11.00		X		O						X
10.75			O							

Tabella 10. Risultati delle prove sulla ruota pallinata S230 Medio A + Z150 2 bar

F_{max} [kN]	1	2	3	4	5	6	7	8	9
11.50	X						X		X
11.25		X		X		O		O	
11.00			O		O				

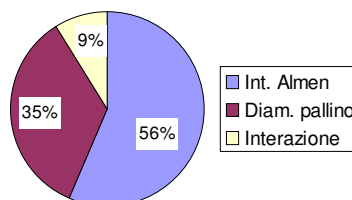
quello per la ruota con pallinatura S110 Basso A non tribofinita, limiti determinati nelle sezioni precedenti.

Come illustrato precedentemente, uno degli obiettivi di questa parte della campagna sperimentale era quello di valutare l'effetto della doppia pallinatura sulle prestazioni a fatica. Come risulta dall'istogramma, i limiti determinati sono tra loro quasi uguali (1247 MPa per S230 Medio A, 1253 MPa per il componente sottoposto al trattamento aggiuntivo), con un aumento di appena 6 MPa (0,5%). In tal caso è immediato concludere che la doppia pallinatura ha un impatto trascurabile sul limite dinamico. Tale dato è in parziale accordo con [1], che aveva osservato come una doppia pallinatura comportasse un debole aumento della resistenza a fatica. Osservando le bande di errore, si nota tuttavia che l'incertezza sul risultato è molto minore nel caso della ruota con doppia pallinatura. Questo dato risulta evidente se si considerano i risultati delle prove di fatica: nelle Tab. 9-10 sono riportate le successioni delle rotture e non rotture [27].

Confrontando la deviazione standard, si osserva che nel primo caso essa ammonta a 35 MPa, diminuiti a 21 MPa nel caso del trattamento aggiuntivo, con un decremento pari al 40%. In questo senso la doppia pallinatura risulta avere un effetto simile a quello già osservato per la tribofinitura: ha un influsso trascurabile sul limite di fatica, ma riduce sensibilmente la dispersione dei risultati e la deviazione standard. E' bene precisare che tali responsi sono stati ottenuti su ruote dentate in acciaio ad elevata resistenza: non si può escludere che miglioramenti più sensibili si possano ottenere su componenti in un acciaio con caratteristiche meno spinte.

Tabella 11. Analisi della varianza sui limiti a fatica e decomposizione della somma di quadrati totale

	SSQ	G. di L.	MSQ	F _{calc.}	p-v. [%]
Int. Almen	4485	2	2242	2.88	6.73
Diam. pallino	2764	2	1382	1.77	18.20
Interazione	696	1	696	0.89	35.00
Errore		42	Stimata: 779		
Totale	7944	47			



Dall'analisi quantitativa dei risultati di Fig. 12, si evince che, grazie al trattamento di pallinatura, si possono ottenere miglioramenti sul limite di fatica fra il 19 ed il 31%, questo campo è compatibile con quanto rilevato anche da alti autori, come [2-3,14]. I dati raccolti sono stati processati in un'analisi ANOVA a due fattori, con calcolo dell'errore sulla base della media pesata delle varianze in accordo con [3,29,37]. L'analisi ha mostrato come l'intensità Almen sia significativa alla soglia del 7% e il diametro del pallino a quella del 18%. Il parametro più significativo risulta quindi quello dell'intensità Almen: questo rende ragione dell'approccio classico per definire l'intensità della pallinatura: quest'ultima ed il conseguente beneficio sulla resistenza a fatica vengono correlate al parametro dell'intensità Almen, che a sua volta assomma tanti fattori che intervengono nel processo [5,45].

Tale interpretazione rende ragione anche del modello proposto da Benedetti, che in [2] correla il limite di fatica al picco di tensione residua e quest'ultimo all'intensità Almen. Tuttavia il diagramma di Tab. 11 mostra come anche il diametro del pallino intervenga per il 35% sul valore della varianza complessiva: come rilevato in [6] anche il diametro del pallino interviene a determinare il profilo di tensione residua e di riflesso ha un influsso sulla resistenza a fatica. Si ha inoltre un piccolo contributo non significativo dovuto all'interazione tra i due fattori. Il fatto che il p-value risulti maggiore del 5% indica che nessuno dei due fattori è altamente significativo sul limite di fatica. Questo a conferma del fatto che si è molto vicini al limite strutturale dei benefici conferiti dal trattamento, sia per le osservazioni prima riportate sia per l'impiego su un materiale ad elevate prestazioni, già incrementate anche dai trattamenti termici di cementazione e tempra, che, ancor prima della pallinatura, introducono delle tensioni residue nel pezzo.

Progettare a fatica. Il modello più semplice è quello di Benedetti che in [2] evidenzia un legame lineare fra limite di fatica e picco di tensione residua. Tale modello viene utilizzato sia per i componenti pallinati che per quelli non pallinati, la cui distribuzione residua è legata alla storia tecnologica. In tale lavoro è evidenziato tuttavia come i risultati per i componenti pallinati e per gli stessi non trattati si collochino su rette interpolanti ben distinte, inoltre in quest'ultimo caso non viene considerata la tensione residua massima, ma quella al di sotto dello strato ossidato superficiale, dove si ritiene più probabile che possa partire la frattura. Nel presente caso si è ricercato il legame fra i limiti trovati sperimentalmente e i massimi di tensione residua: nel caso dei componenti pallinati si sono considerati i valori di picco: nell'unico caso di componente non pallinato si è fatto riferimento al valore più elevato sulla superficie (Fig. 9). In accordo con [1,8] è infatti sullo strato più superficiale che in componenti non pallinati si innesca il fenomeno di rottura a fatica.

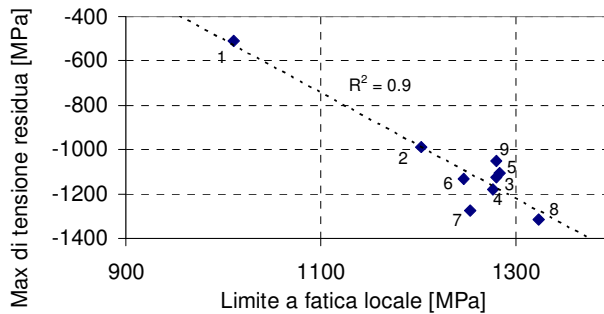


Figura 13. Legame tra i limiti di fatica e i massimi di tensione residua

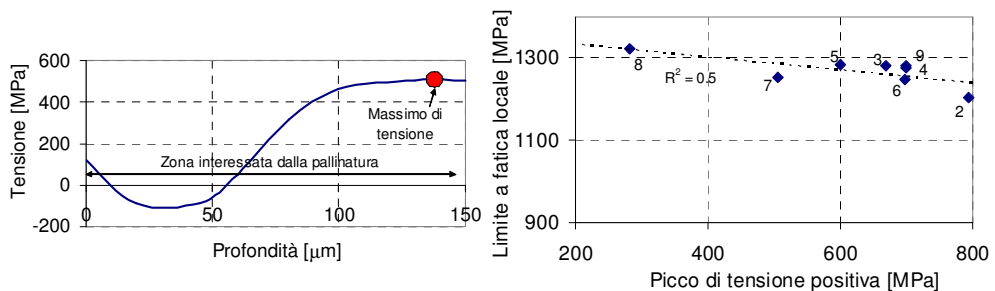


Figura 14. Tipica distribuzione di tensione al piede e legame tra i limiti di fatica e i massimi di tensione al di sotto della superficie

Il modello proposto, per quanto semplice, si presenta piuttosto realistico: tutti i risultati sperimentali (Fig. 13), anche quello per le ruote non pallinate, risultano ben interpolati da una linea retta di regressione, come testimoniato dall'alto valore di R^2 (coefficiente di correlazione lineare). I numeri si riferiscono alla legenda di Fig. 12.

Il successivo modello analizzato è stato quello suggerito da Wang in [8]. Questi osserva come le tensioni residue possano essere interpretate come un incremento od un decremento degli sforzi da carico esterno. Evidenzia inoltre come le tensioni residue da pallinatura possano portare il punto di innesco dallo strato superficiale ad uno sub-superficiale. Quest'ultimo si può collocare nella zona immediatamente al di sotto della superficie, dove sono presenti le tensioni residue di compressione, o ancora più in profondità, dove le tensioni residue divengono di trazione. Risulta ragionevole riferire il limite di fatica alla massima tensione positiva al di sotto della superficie. Come risulta dalla Fig. 14 (che riporta la distribuzione totale di tensione per la ruota pallinata 230 Medio A + Z150, 2 bar), questo massimo è situato in genere poco al di sotto della superficie, entro o comunque ai limiti della zona interessata dalla pallinatura. Ciò appare dovuto essenzialmente alla piccola dimensione della sezione trasversale dei denti (8 mm) ed all'elevata concentrazione di tensione al piede. La Fig. 14 riporta i limiti di fatica sperimentali rispetto alle massime tensioni positive all'interno. In questo caso non si è considerato il componente non pallinato, visto che, come già ravvisato, la rottura a fatica avviene per esso secondo un meccanismo diverso.

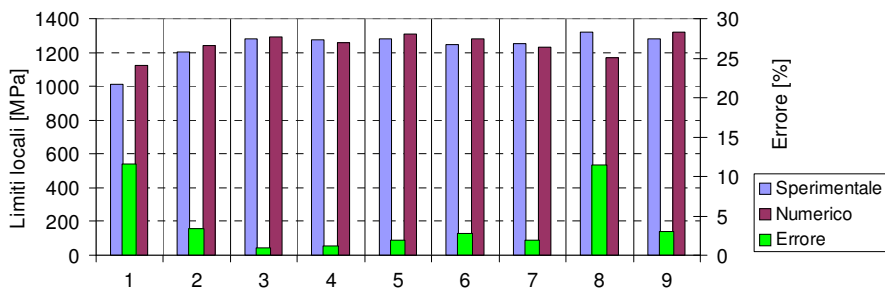


Figura 15. Metodo di Eichlseder: confronto fra risultati sperimentali e numerici e calcolo degli errori percentuali

In tale caso l'interpolazione, pur sempre abbastanza buona, non è però ottimale, come indicato dal valore di 0,5 per il coefficiente R^2 . E' comunque significativo notare come i punti 4 e 9 (rispettivamente S110 Medio A e S110 Basso A senza tribofinitura, cfr. legenda di Fig. 12) siano fra loro quasi coincidenti: alla medesima tensione di trazione massima corrisponde il medesimo limite a fatica.

Secondo lo schema già seguito a proposito della simulazione dell'intera curva di Wöhler, il modello di Eichlseder [22] è stato utilizzato per effettuare previsioni numeriche dei limiti di fatica. La distribuzione totale è stata quindi ottenuta, sommando alle tensioni simulate quelle residue dovute alla pallinatura, Fig. 9. Il modello proposto prevede di considerare i punti di maggiore criticità sempre in superficie: è quindi sullo strato superficiale che si sono calcolati (con l'applicazione di (2)) i gradienti di tensione relativi per tutti i trattamenti considerati. Tramite la (3) si è quindi determinata la resistenza a fatica locale, tenendo conto del profilo di tensioni residue sia per la stima della distribuzione di tensione totale, sia per aggiustare il limite trovato in funzione della tensione media [32,42-44]. I risultati sono mostrati nella Fig. 15, che confronta i rilievi sperimentali e le previsioni numeriche per i nove casi in studio, evidenziando anche gli errori. Il modello proposto è stato applicato anche al caso della ruota non pallinata, valutando le sole tensioni residue da trattamento termico. Come si può notare dalla Fig. 15, gli errori, sia in eccesso che in difetto, sono abbastanza contenuti ed accettabili, l'errore medio è di circa il 4%, mentre il massimo ammonta a poco più dell'11% ed è stato rilevato nel caso del componente pallinato.

L'analisi effettuata può essere posta a confronto con quella operata da Benedetti in [2]: Benedetti utilizza per le sue stime una formulazione derivante anch'essa dalla scuola tedesca, in particolare dalle esperienze di Lang [46]. Pure tale modello correla la resistenza a fatica (a sollecitazione alternata) al concetto di gradiente (legame di tipo lineare), ma fa riferimento al gradiente puro della distribuzione e non al suo valore normalizzato rispetto alla tensione puntuale. Tale modello inoltre effettua previsioni di fatica, a partire da dati statici, quali la rugosità superficiale e la durezza Vickers, non sempre adeguati per la valutazione del fenomeno della fatica. Inoltre esso non è applicabile a componenti contenenti tensioni residue. Le previsioni numeriche sono state attuate, considerando inizialmente solo il carico esterno (e non le tensioni residue), correggendo poi i limiti trovati sulla base della tensione media relazionata al solo valore di picco delle tensioni residue. La semplificazione del modello ha comportato errori di sottostima di circa il 15%.

In [2] sono riportate numerose ragioni, che spiegano in dettaglio gli scostamenti trovati: si può al proposito aggiungere che un'ulteriore causa potrebbe essere legata alla scarsa robustezza del modello previsionale di Lang. Con riferimento ai metodi basati sul gradiente, va rilevato che in presenza di valori elevati di questo (distribuzioni tensionali molto ripide) è spesso difficile determinare il gradiente massimo per via simulativa. Questo richiederebbe un affinamento eccessivo e non praticabile del modello agli elementi finiti: è quindi essenziale che gli eventuali errori a monte si ripercuotano il meno possibile sul calcolo del limite di fatica. Il modello lineare di Lang comporta che un eventuale errore sulla stima del gradiente è trasmesso inalterato sulla valutazione della resistenza a fatica; il modello di Eichlseder invece è non lineare (il termine K_D vale 0,3) ed ha un andamento simile a quello di una curva logaritmica, per cui risulta molto stabile per gradienti sufficientemente elevati: nel caso della presente ricerca un errore in difetto del 30% sulla valutazione del gradiente di tensione relativo avrebbe comportato scostamenti fra l'1 ed il 3% rispetto alle previsioni del limite di fatica sopra mostrate.

I risultati della sperimentazione sono stati interpretati anche alla luce del modello proposto da Taylor nella sua Teoria delle distanze critiche (Theory of critical distances, TCD). Tale impostazione metodologica è stata oggetto di numerosi articoli scientifici, svolti anche in collaborazione con ricercatori italiani, ed è stata recentemente raccolta nel testo [23]. Il modello proposto da Taylor, oltre che fornire un'interpretazione del fenomeno della fatica, mette a disposizione uno strumento di previsione del limite strutturale del componente. Il concetto di limite strutturale va considerato in senso lato, potendosi applicare tale approccio non solo al caso di organi sollecitati a fatica, ma anche a quello di corpi caricati staticamente. La teoria sviluppata poggia sull'assunto che, nel caso della fatica, la rottura di un pezzo si verifica qualora la variazione di tensione locale in una localizzazione posta ad una "distanza critica" dal punto di concentrazione raggiunga un "valore critico". È questo il cosiddetto Metodo del Punto (Point Method, PM), mentre il metodo della linea prende in considerazione non lo stato tensionale in un solo punto, ma quello mediato su una linea, avente lunghezza legata alla distanza critica. La valutazione del comportamento a fatica si risolve quindi nel calcolo della distanza critica, nella valutazione della variazione di tensione nella posizione da tale distanza critica individuata e nel confronto rispetto al valore critico, corrispondente al limite a fatica del materiale.

Rispetto alla teoria di Eichlseder ed in generale alle altre teorie del gradiente si possono ravvisare i seguenti punti in comune. Ambedue i metodi fanno affidamento ad analisi agli elementi finiti. Entrambi gli approcci richiedono dati di caratterizzazione del materiale, ricavabili da prove sperimentali su provini. Nel caso del metodo TCD sono richiesti i parametri del ΔK_{th} (valore di soglia per la variazione del fattore di intensità degli sforzi) e del limite di fatica. Ambedue i metodi richiedono quindi almeno due parametri e l'analisi di un fattore di controllo, che è dato dal gradiente relativo per il primo caso e dallo stato tensionale locale alla distanza critica per il secondo. Entrambe le teorie riconoscono come la fatica non sia un fenomeno puramente locale: non è cioè il solo picco massimo di tensione a portare alla rottura del pezzo, ma è la distribuzione tensionale in un volume attiguo. Rispetto a queste basi comuni le due teorie si sono evolute in modo diverso, l'una chiamando in causa il concetto del gradiente di tensione e l'altra quello della distanza critica. Se nel primo caso il concetto di effetto supporto è chiaramente messo in luce (riferendosi all'effetto di sostegno delle fibre più interne normalmente meno caricate, a seguito della microplasticizzazione degli strati dove si riscontrano i picchi di tensione), nel secondo è evitato con una definizione di criticità ad una certa distanza. Entrambe le teorie

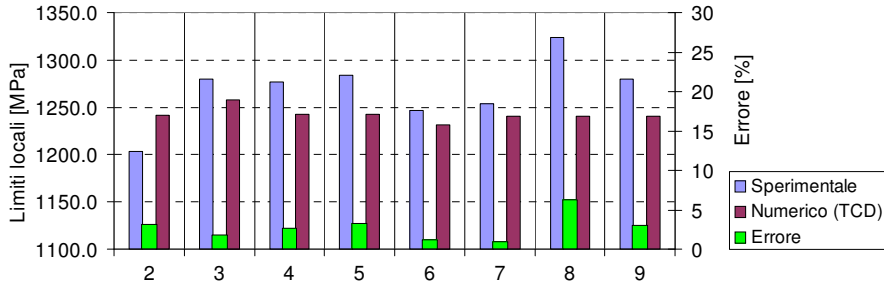


Figura 16. Metodo di Taylor: confronto fra risultati sperimentali e numerici e calcolo degli errori percentuali

hanno il vantaggio di essere coerenti con gli approcci classici nel caso di corpi desanvenanziani, e quello di essere generalizzabili all'applicazione su corpi tozzi. Gli svantaggi possono essere rappresentati dalla difficoltà nel calcolo esatto del gradiente relativo per quanto concerne il metodo di Eichlseder, pur restando valide le considerazioni sulla robustezza di quest'ultimo, e dalla difficile reperibilità o misura dei dati di ΔK_{th} per quanto riguarda il metodo di Taylor.

Dopo la presente analisi critica, si è applicato il metodo di Taylor al caso in esame, valutandone l'attendibilità dei risultati, pur in assenza di dati precisi di ingresso. Non disponendo di un dato certo sul ΔK_{th} , lo si è assunto pari a $6 \text{ MPa m}^{0.5}$, sulla base di dati riportati in letteratura e anche dallo stesso Taylor [23]. A rigore tale dato varrebbe solo per il rapporto di carico R pari a 0: in [23,47] si rimarca il fatto che, al variare di R , variano anche il limite di fatica ed il valore di soglia ΔK_{th} (e di conseguenza pure la distanza critica), tuttavia in [23] è precisato come il metodo proposto sia piuttosto robusto anche su un ampio campo di variazione di R . Per semplicità ci si è limitati all'applicazione del Metodo del Punto, ricavando la lunghezza critica L come segue (Eqn. (7)):

$$L = \frac{1}{\pi} \left(\frac{\Delta K_{th}}{\sigma_D} \right)^2 = 8 \mu m \quad (7)$$

Per σ_D si è inteso il limite di fatica del materiale, ricavato da prove su provini lisci, pari a 1175 MPa. Il risultato ottenuto appare incoraggiante: la lunghezza critica ha infatti un'estensione decisamente esigua ed è ampiamente compresa entro lo strato pallinato. Questo dimostra che il metodo di Taylor può essere validamente applicato pure al caso di organi pallinati. Successivamente si sono valutate le distribuzioni tensionali compressive (dovute al carico esterno e alla tensione residua) per tutti i componenti diversamente pallinati precedentemente considerati, utilizzando i modelli agli elementi finiti già impiegati per le precedenti elaborazioni. La variazione di tensione in un ciclo alla distanza critica è stata quindi paragonata e rapportata al limite di fatica di 1175 MPa: applicando tale rapporto di scala si sono ricavate le stime sui limiti locali. Questi sono stati confrontati con quelli sperimentali, come mostrato in Fig. 16. Si evidenzia un buon accordo fra dati

sperimentali e previsioni numeriche, con uno scostamento massimo del 6 % ed uno minimo del 3%.

4. CONCLUSIONI

Il presente lavoro ha riguardato lo studio del comportamento a fatica di ruote dentate ad elevate prestazioni per uso nel settore della competizione automobilistica. Tale tematica è stata affrontata su più livelli: tramite rilievo sperimentale delle tensioni residue e della resistenza a fatica, tramite l'effettuazione di elaborazioni DOE e attraverso l'impiego di modelli per valutare la possibilità di un progetto a fatica sulla base di simulazioni. I principali risultati possono essere riassunti come segue.

- Un trattamento di pallinatura abbastanza comune, non eccessivamente spinto, comporta un aumento del 19% della resistenza.
- La tribofinitura ha un effetto trascurabile sulla resistenza a fatica di componenti ad elevate prestazioni come quelli considerati. Ha invece un ruolo importante nella riduzione della dispersione sperimentale e della deviazione standard (-57%). La tribofinitura non va inoltre ad intaccare la distribuzione residua dovuta alla pallinatura.
- In riferimento ad uno specifico e diffuso trattamento di pallinatura si è prima determinata sperimentalmente la curva S-N, successivamente lo stesso problema è stato affrontato per simulazione. Sulla base del modello del gradiente relativo di Eichlseder e combinando i risultati di analisi strutturali e di impatto FEM, si è ricavata una curva simulata, riscontrando un buon accordo (scostamenti minori del 10%) per quanto riguarda il valore del limite e quello della pendenza del tratto inclinato.
- Si è ricercata un'ottimizzazione del trattamento di pallinatura, organizzando in un piano fattoriale diverse combinazioni del diametro del pallino e dell'intensità Almen. Lo studio diffrattometrico delle tensioni residue è stato accompagnato da un'analisi statistica, che ha confermato come la profondità del picco sia legata principalmente al diametro del pallino e il valore del picco all'intensità Almen, tale legame è però fortemente non lineare. La tendenza alla saturazione dimostra il limite strutturale della pallinatura: non si hanno apprezzabili vantaggi ad incrementare fortemente i parametri in gioco. Si è mostrato come si possa raggiungere un incremento del limite di fatica fino al 31% rispetto all'organo non pallinato: un'analisi DOE mostra come sia l'intensità Almen il parametro più influente.
- Si è valutato anche l'influsso della doppia pallinatura: un trattamento aggiuntivo con pallini ceramici comporta un forte incremento sulla tensione di superficie, e rende l'andamento delle tensioni residue sostanzialmente costante alle deboli profondità. Non si hanno forti variazioni sul valore di picco: coerentemente con questo l'influenza sul limite di fatica è poco significativa. La doppia pallinatura ha un ruolo molto simile a quello della tribofinitura nell'aumentare sensibilmente la ripetibilità dei risultati, riducendo la deviazione standard (-40%).
- I risultati ottenuti sono stati interpretati con l'uso di modelli simulativi: essi risultano in buon accordo con la formulazione di Benedetti: le resistenze a fatica sono legate linearmente rispetto ai valori massimi le tensioni residue. Rispetto al succitato modello si è osservato come tale legame valga anche per il componente non pallinato, per il quale le tensioni residue derivano dai trattamenti termici da esso subiti.
- Si è effettuata un'analisi comparativa sui metodi di Eichlseder gradiente relativo e di Taylor (Teoria delle distanze critiche), evidenziandone i lati in comune, i punti di forza

e gli svantaggi. Successivamente sono stati sviluppati modelli di previsione basati su entrambe le teorie, ivi applicati alla casistica in esame dei corpi pallinati. I risultati preliminari appaiono positivi, anche alla luce di incertezze di dati sul materiale: lo scostamento fra dati sperimentali e previsioni numeriche non supera l'11% (con un valore medio del 4%) nel caso del modello di Eichlseder ed il 6% (con un valore medio attorno al 3%) nel caso del modello di Taylor.

- Analisi di sensitività sono state effettuate sul modello del gradiente relativo: esso si è dimostrato apprezzabilmente robusto: sottostime del valore del gradiente, in cui si può facilmente incorrere nelle simulazioni numeriche della distribuzione di tensione, non si ripercuotono gravemente sui risultati finali.

BIBLIOGRAFIA

- [1] Shaw B.A., Aylott C., O'Hara P., Brimble K., 2003, "The role of residual stress on the fatigue strength of high performance gearing", *International Journal of Fatigue*, **25**, pp. 1279-1283.
- [2] Benedetti M., Fontanari V., Höhn B.R., Oster P., Tobie T., 2002, "Influence of shot peening on bending tooth fatigue limit of case hardened gears", *International Journal of Fatigue*, **24**, pp. 1127-1136.
- [3] Croccolo D., Cristofolini L., Bandini M., Freddi A., 2002, "Fatigue strength of shot peened nitrided steel: optimization of process parameters by means of design of experiment", *Fatigue and Fracture of Engineering Materials and Structures*, **25**, pp. 695-707.
- [4] Mann JY., *Fatigue of materials—an introductory text*, Melbourne University Press, 1967.
- [5] Guagliano M., Vergani L., 2004, "An approach for prediction of fatigue strength of shot peened components", *Engineering Fracture Mechanics*, **71**, pp. 501-512.
- [6] Herzog R., Zinn W., Sholtes B., Wohlfahrt H., 1996, "The significance of Almen intensity for the generation of shot peening residual stresses", In Proceedings of the VI International Conference on Shot-Peening, Vol. **1**, pp. 270-281.
- [7] Kobayashi M., Matsui T., Murakami Y., 1998, "Mechanism of creation of compressive residual stress by shot peening", *International Journal of Fatigue*, **20**, pp. 351-357.
- [8] Wang S., Li Y., Yao M., Wang R., 1998, "Fatigue limits of shot-peened metals", *Journal of Materials Processing Technology*, **73**, pp. 57-63.
- [9] Jiang X.P., Man C.S., Shepard M.J., Zhai T., 2007, "Effects of shot-peening and re-shot-peening on four-point bend fatigue behaviour of Ti-6Al-4V", *Materials Science and Engineering A*, **468-470**, pp. 137-143.
- [10] Colombo C., Guagliano M., Vergani L., 2005, "Pallinatura di acciai nitrurati: analisi sperimentale e numerica", In Proceedings of the XXXIV AIAS National Conference, Vol. **1**.
- [11] Fathallah R., Laamouri A., Sidhom H., Braham C., 2004, "High cycle fatigue behavior prediction of shot-peened parts", *International Journal of Fatigue*, **26**, pp. 1053-1067.
- [12] Wagner L., 1999, "Mechanical surface treatments on titanium, aluminium and magnesium alloys", *Materials Science and Engineering A*, **263**, pp. 210-216.
- [13] Song P.S., Wen C.C., 1999, "Crack closure and crack growth behavior in shot peened fatigued specimen", *Engineering Fracture Mechanics*, **63**, pp. 295-304.
- [14] Ohsawa M., Yonemura T., 1984, "Improvement of hardened surface by shot peening", In

Proceedings of the II International Conference on Shot Peening (ICSP2), pp. 147-158.

[15] Benedetti M., Azanza Ricardo C.L., Santus C., Fontanari V., 2007, "Studio dell'effetto della pallinatura sul comportamento a fatica della lega Al-7075-T651", In Proceedings of the XXXVI AIAS National Conference, Vol. 1.

[16] Winkelmann L., Michaud M., Sroka G., Swiglo A.A., 2002, "Impact of Isotropic Superfinishing on Contact and Bending Fatigue of Carburized Steel", In Proceedings of the SAE International Off-Highway Congress.

[17] Ishibashi A., Ezoe S., Tanaka S., 1984, *Mirror Finishing of Tooth Surfaces Using a Trial Gear Grinder With Cubic-Born-Nitride Wheel*, ASME Publication 84-DET-153, ASME, New York, NY.

[18] Krantz T.L., Alanou M.P., Evans H.P., Snidle R.W., 2001, "Surface Fatigue Lives of Case-Carburized Gears With an Improved Surface Finish", *Journal of Tribology*, **123**, 709-716.

[19] Breen D.H., Wene E.M., 1979, "Fatigue in Machines and Structures – Ground Vehicles", In Proceedings of the American Society for Metals, Metals Park, p. 72.

[20] Fuchs H.O., 1986, *Mechanical engineer's handbook*, Wiley, New York.

[21] Siebel E., Pfender M., 1947, "Weiterentwicklung der Festigkeitsberechnungen bei Wechselbeanspruchung", *Stahl und Eisen*, **66/67**, pp. 318-321.

[22] Eichlseder W., 2002, "Fatigue analysis by local stress concept based on finite element results", *Computers and Structures*, **80**, pp. 2109-2113.

[23] Taylor D., 2007, *The Theory of Critical Distances - A new perspective in fracture mechanics*, Elsevier Science B.V..

[24] Pariente I.F., Guagliano M., 2007, "About the role of residual stresses and surface work hardening on fatigue ΔK_{th} of a nitrided and shot peened low-alloy steel", *Surface & Coatings Technology*, in press, doi: 10.1016/j.surfcoat.2007.11.015.

[25] Guagliano M., Riva E., Guidetti M., 2002, "Contact fatigue failure analysis of shot-peened gears", *Engineering Failure Analysis*, **9**, pp. 147-158.

[26] Peyrac C., Flavenot J.F., Convert F., 2000, "Combining case hardening and shot peening for gear steels: influence on residual stresses", *Materials Science Forum*, **347**, pp. 435-440.

[27] Olmi G., Freddi A., 2006, "Shot peening technique for fatigue improvement of high strength steel gears", In Proceedings of the 5th ICCSM.

[28] Comandini M., Olmi G., Freddi A., 2007, "Fatigue performance of shot-peened gears investigated by experimental and numerical methods", *Transactions of Famena*, **31-2**, pp. 1-10.

[29] Dixon W.J., Massey F.J.Jr., 1983, *Introduction to Statistical Analysis*, McGraw-Hill, U.S..

[30] Benedetti M., Bortolamedi T., Fontanari V., Frendo F., 2004, "Bending fatigue behaviour of differently shot peened Al 6082 T5 alloy", *International Journal of Fatigue*, **26**, pp. 889-897.

[31] Moore M.G., 1958, *SAE Transactions*, **66**, p. 340.

[32] Bertini L., Fontanari V., 1999, "Fatigue behaviour of induction hardened notched components", *International Journal of Fatigue*, **21** (6), pp. 611-617.

[33] Meguid S.A., Hammond D.W., 1988, "Fatigue fracture and residual stress relaxation in shot peened components". In Proceedings of the Second International Conference on Residual Stresses, pp. 797-802.

[34] Cao W., 1989, "Cyclic softening of steels and residual stress relaxation of shot-blasting in

fatigue”, Ph.D. thesis, ENSAM Paris, France.

[35] Zhuang W.Z., Halford G.R., 2001, “Investigation of residual stress relaxation under cyclic load”, *International Journal of Fatigue*, **23**, pp. 31-37.

[36] Comandini M., Olmi G., Freddi A., 2006, “Comparison between theoretical and experimental evaluation of the fatigue life of shot-peened gears”, In Proceedings of the 23rd DAS Symposium, pp. 163-164

[37] Berger P.D., Maurer R.E., 2002, *Experimental Design*, Boston, USA.

[38] ISO 12107, 2003, *Metallic materials - Fatigue testing - Statistical planning and analysis of data*.

[39] Guagliano M., 2001, “Relating the Almen intensity to residual stresses induced by shot peening: a numerical approach”, *Journal of Materials Processing Technology*, **110**, pp. 277-286.

[40] Benedetti M., Fontanari, V., Oster P., Tobie T., 2001, “Influence of residual stresses on fatigue behavior of surface treated gears”, In Proceedings of the XXX AIAS National Conference, pp. 263-272.

[41] Reed-Hill R.E., Abbaschian R., 1992, *Physical Metallurgy Principles*, Publishing Company, Boston, MA, USA.

[42] Starker P., Macherauch E., 1983, “Shot peening and fatigue endurance”, *Z. Werkstofftechnik*, **14**, pp. 109-115.

[43] Leitner H., Eichlseder W., Godor I., Waggermayer M., Hinteregger C., *Increased Fatigue Limit of Gear Wheels by a Combination of Case Hardening and Shot Peening*, Internal Report of the MontanUniversität Leoben, Austria.

[44] Eichlseder W., Leitner H., *Fatigue life prediction based on gradient S-N curves*, Internal Report of the MontanUniversität Leoben, Austria.

[45] Guagliano M., Vergani L., Bandini M., Gili F., 1999, “An approach to relate the shot-peening parameters to the induced residual stresses”, In Proceedings of the ICSP7, pp. 274-282

[46] Lang O.R., 1979, “Dimensioning of complicated steel machine parts in the whole range of load cycle and endurance”, *Z. Werkstofftechnik*, **10**, pp. 24-29.

[47] Atzori B., Meneghetti G., Susmel L., 2005, “Material fatigue properties for assessing mechanical components weakened by notches and defects”, *Fatigue and Fracture of Engineering Materials and Structures*, **28**, pp. 83-97

SULLA RESISTENZA RESIDUA A TORSIONE DOPO IMPATTO DI TUBI IN MATERIALE COMPOSITO

Giangiaco Minak

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Aeronautiche, Nucleari e di Metallurgia

Alma Mater Studiorum – Università di Bologna, Italia

E-mail: giangiaco.minak@unibo.it

Daniele Ghelli

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Aeronautiche, Nucleari e di Metallurgia

Alma Mater Studiorum – Università di Bologna, Italia

E-mail: daniele.ghelli@mail.ing.unibo.it

Andrea Zucchelli

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Aeronautiche, Nucleari e di Metallurgia

Alma Mater Studiorum – Università di Bologna, Italia

E-mail: andrea.zucchelli@mail.ing.unibo.it

Abstract. *Le strutture tubolari in materiale composito a fibra lunga sono recentemente diventate di notevole interesse industriale per le possibilità offerte dalle elevate proprietà meccaniche e dal peso limitato anche se ad un costo superiore a quello dei materiali convenzionali. Questo si traduce in migliori prestazioni nel campo sportivo, meno sensibile ai costi, ma anche in settori industriali dove l'incremento di velocità reso possibile dalla riduzione della massa e dall'aumento del valore delle frequenze proprie dovute alla maggiore rigidità sia fondamentale. Un problema che si deve porre il progettista è di valutare la capacità di mantenere queste proprietà nel tempo durante l'uso. In questo lavoro viene misurata sperimentalmente la resistenza residua a torsione di tubi in composito soggetti ad impatto laterale e viene confrontata con la resistenza dei tubi intatti al fine di elaborare criteri progettuali che consentano di sfruttare al meglio il materiale.*

Keywords: *Impatto, Torsione, Resistenza Residua, Materiale Composito*

1. INTRODUZIONE

I materiali compositi a matrice polimerica rinforzati mediante fibra di carbonio, denominati sovente CFRP dall'acronimo in lingua Inglese, sono caratterizzati da proprietà meccaniche specifiche superiori a quelle di gran parte dei materiali tradizionali e questo consente di ottenere componenti di resistenza elevata e peso contenuto.

Nella maggior parte dei casi questi materiali si presentano in forma di laminati ottenuti mediante la polimerizzazione di lamine sovrapposte, ciascuna con una direzione preferenziale per la resistenza meccanica, vale a dire quella lungo la quale sono disposte le fibre. Questa caratteristica, da un lato consente di ottenere componenti con resistenza maggiore in determinate direzioni, la cui scelta dipende dalle azioni esterne agenti sul componente, dall'altro però fa sì che la resistenza a carichi trasversali al piano del laminato sia limitata a causa della tendenza alla delaminazione. Una volta delaminato, il materiale non riesce a sostenere tensioni di compressione a causa dell'insorgere dei fenomeni di instabilità elastica a livello di lamina o di singolo trefolo di fibre.

La problematica della resistenza residua dei laminati piani, che si presenta soprattutto nel caso di carichi impulsivi, è stata studiata in maniera estesa in letteratura e anche dagli autori [1, 2]

Recentemente i CFRP sono stati applicati in campo industriale per strutture tubolari soggette a torsione quali alberi di verricelli per barche a vela o alberi di macchine per l'industria cartiera.

Nella letteratura scientifica questi tubi in composito sono stati studiati per lo più per quanto riguarda la resistenza e la resistenza residua dopo impatto quando soggetti a pressione interna o esterna [3,4]. In particolare nel caso di sollecitazioni torsionali il problema che è stato più studiato è quello dell'instabilità dell'equilibrio elastico globale in strutture a parete sottile [5].

Nel presente lavoro viene effettuata la caratterizzazione di tubi in CFRP, di spessore sufficientemente elevato in rapporto al diametro da non risentire dell'instabilità, soggetti a torsione ed impatto trasversale.

Per differenti sequenze di laminazione sono valutate in particolare: la resistenza a torsione, la resistenza residua a torsione dopo impatto e l'effetto della presenza di un precarico torsionale al momento dell'impatto sulla resistenza residua.

2. MATERIALI E METODI

Quattro serie di tubi in materiali composito lunghi 200mm di diametro interno 50mm e 2,55 mm di spessore sono stati prodotti mediante polimerizzazione in autoclave di stuoie unidirezionali di fibra di carbonio T300 preimpregnate con resina epossidica (fibra 60% in volume). Le sequenze di laminazione utilizzate, mostrate in tabella1, sono due incrociate con fibre disposte a 45° rispetto all'asse del tubo, al fine di massimizzare la resistenza torsionale, e due quasi isotrope. Nella pratica non si può utilizzare la prima configurazione perché molto vulnerabile a carichi assiali o flessione, mentre la seconda presenta un numero eccessivo di lamine con angolazione diversa da 45°. I risultati ottenuti potranno essere quindi considerati come due limiti, superiore ed inferiore, delle prestazioni torsionali di un componente di forma tubolare.

Ai campioni così ottenuti sono stati incollati degli appositi afferraggi in acciaio, studiati [6] per minimizzare la concentrazione di tensioni ed evitare la rottura in corrispondenza delle zone vincolate.

Differenti combinazioni di prove sono state eseguite mediante una macchina torsionale servoidraulica (si veda la figura 1) da 50 kNm equipaggiata con un pendolo strumentato con una cella di carico piezoelettrica (PCB Piezotronics, Model 208A35) dotata di testa emisferica da 12.7 mm di diametro. La geometria del sistema era fatta in modo che la testa

Tabella 1. Sequenze di laminazione

Tipo	Sequenza
CPS	[45/-45] _{4s}
CPNS	[45/-45] ₈
QIS	[02/45/-45/45/-45/90] _{2s}
QINS	[0 ₂ /45/-45/45/-45/90 ₄ /45/-45/45/-45/0 ₂]

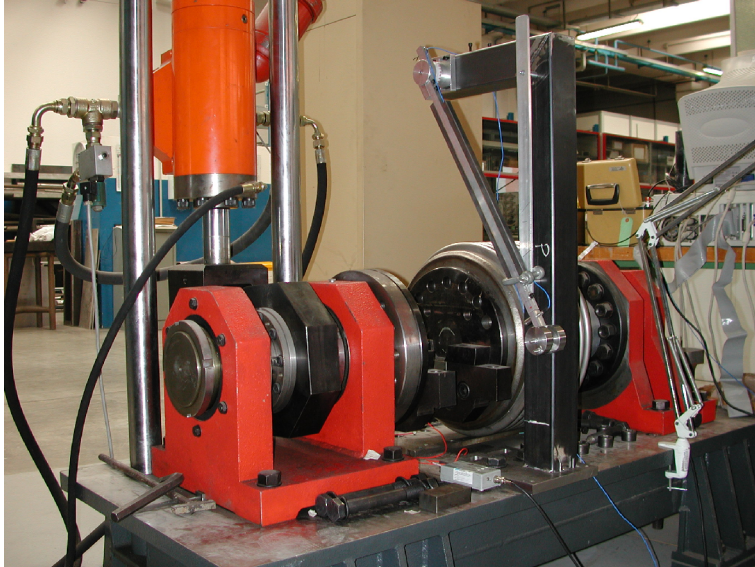


Figura 1. macchina servoidraulica e pendolo strumentato

della cella colpisse i tubi nel centro in maniera ortogonale alle direttrici. L'energia d'impatto stimata era di 7 J, corrispondente come ordine di grandezza a quella di un piccolo attrezzo che cada dall'altezza di circa un metro, mentre la velocità era di 3.5 m/s.

In particolare le prove effettuate con questo sistema sono state di torsione fino a rottura di provini intatti, prove di impatto seguite immediatamente dalla torsione, prove di impatto con precarico torsionale, imposto in controllo di posizione, seguite anche esse da torsione a rottura.

Il precarico torsionale è stato imposto a due livelli: 65% e 130% della resistenza residua a torsione dei provini intatti.

Al livello più basso di precarico si può essere ragionevolmente sicuri che i provini fossero ancora privi di danneggiamento poiché non avevano prodotto alcuna emissione acustica (usualmente provocata dal rilascio di energia elastica dovuto al danneggiamento) che veniva acquisita mediante un sistema di registrazione PAC-DISP4 della Physical Acoustic Corporation equipaggiato con un sensore piezoelettrico di tipo Nano sempre della PAC. Il segnale era preamplificato a 40dB ed era impostato un filtro passa alto a 40dB.

Una particolare funzione (definita nell'eq.1) delle energia acustica emessa cumulata E_a e dell'energia di deformazione E_s (definita nell'eq. 2 dove C è la coppia) chiamata *Sentry* (letteralmente “sentinella”) è stata utilizzata per determinare gli eventi di danno importanti [7].

$$f(\theta) = \ln \frac{E_s(\theta)}{E_a(\theta)} \quad (1)$$

$$E_s(\theta) = \int_0^\theta C \cdot d\theta \quad (2)$$

In corrispondenza di questi eventi è stata calcolata una funzione di danno (eq. 3) basata sul calo della rigidità torsionale.

$$Danno(\theta) = 1 - \frac{(G \cdot J_p)\theta}{(G \cdot J_p)_{iniziale}} \quad (3)$$

3. RISULTATI E DISCUSSIONE

In figura 2 è riportato il comportamento a torsione di quattro campioni con differente sequenza di laminazione. Come ci si poteva aspettare, la rigidità torsionale ed il carico di rottura sono notevolmente più elevati nel caso dei CP, mentre la rotazione a rottura è maggiore per i QI.

Si nota che l'effetto della non simmetria è limitato e si manifesta in maniera evidente solo nei QI oltre la metà del carico di rottura.

In figura 3 si può vedere un esempio della determinazione degli eventi di danno per un provino non impattato attraverso la funzione *Sentry*. Come si può notare, questa funzione è in grado di discriminare eventi avvenuti a partire da carichi molto inferiori a quello di rottura, solitamente riferibili a rotture di matrice o delaminazioni localizzate in corrispondenza delle concentrazioni di tensione dovute agli afferraggi fino agli ultimi istanti corrispondenti alla rottura delle fibre.

Nelle figure 4a e 4b è mostrato l'aspetto delle rotture finali dei provini soggetti a torsione.

I campioni CP mostrano la rottura della lamina esterna, mentre i QI presentano la lamina esterna, che ha le fibre disposte lungo l'asse, completamente distaccata dalla sottostante e con un esteso fenomeno di rottura della matrice lungo la direzione delle fibre (splitting).

Il comportamento ad impatto nelle varie configurazioni, con o senza precarico, è stato estesamente trattato altrove [8, 9].

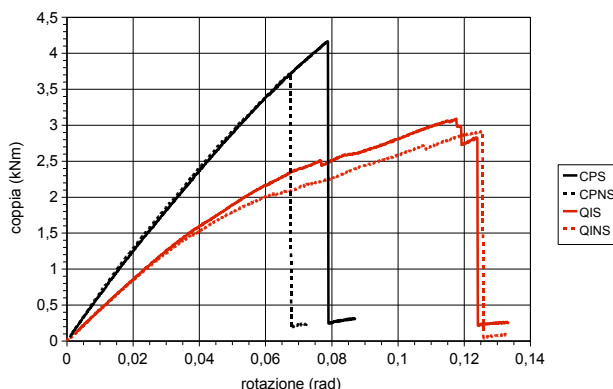


Figura 2. Comportamento a torsione di provini intatti

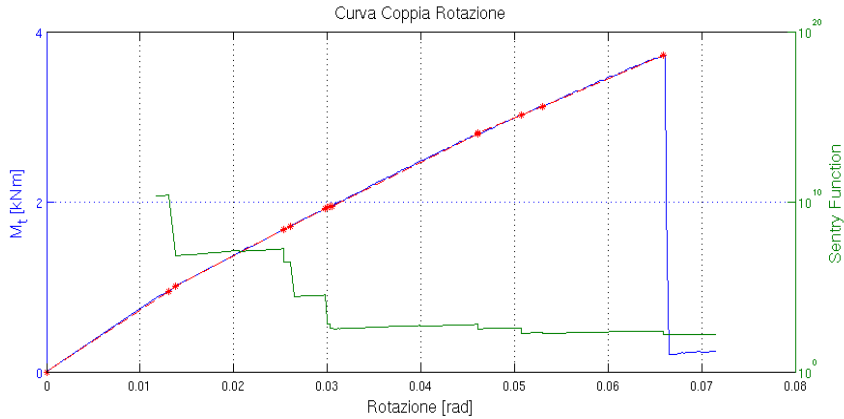


Figura 3. Determinazione degli eventi di danno mediante la funzione Sentry

I risultati riportati in questi lavori indicano che il meccanismo di danneggiamento consiste nella rottura di fibra e matrice nella zona di primo contatto tra impattatore e campione, seguito dalla delaminazione di almeno due coppie di interfacce (la più esterna e la più interna).

Il ruolo del precarico torsionale è quello di fornire energia di deformazione disponibile per i processi di danneggiamento, portando quindi ad un livello di danno finale molto maggiore.

Passando ad investigare il comportamento a torsione dopo l'impatto, dalla figura 5 possiamo notare come un impatto dall'energia tutto sommato limitata possa ridurre sia la capacità di carico sia, in maniera ancora maggiore, la deformazione a rottura. La rigidità dei provini invece rimane per lo più inalterata, confermando la natura essenzialmente legata alla matrice del danno da impatto (quindi rotture di matrice e delaminazione).

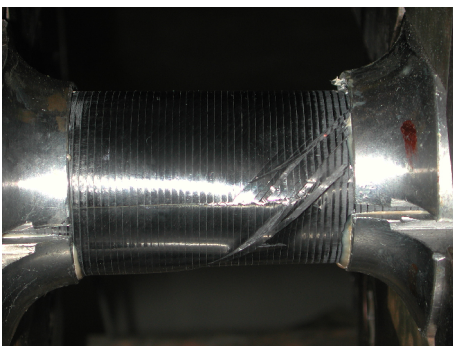


Figura 4a. Rottura a torsione di un campione CP

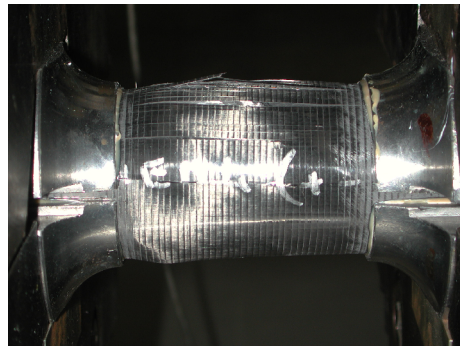


Figura 4b. Rottura a torsione di un campione QI

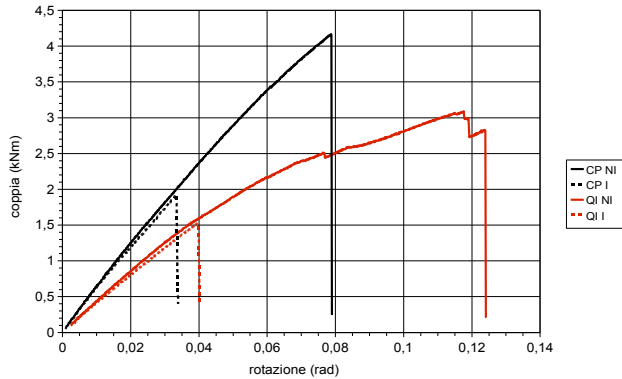


Figura 5. Comportamento a torsione di campioni intatti e danneggiati

Se andiamo a fare un grafico dalla funzione danno come calcolata dalla (3) possiamo osservare (si veda la figura 6) che per i provini QI la curva è divisa essenzialmente in tre zone: la prima in cui il danno cresce linearmente col momento applicato, che è riconducibile al collasso per instabilità locale a compressione dei trefoli di fibre in corrispondenza delle zone delaminate; un plateau in cui il danneggiamento coinvolge la matrice nelle lamine a 0° e 90° provocando limitati cambiamenti nella rigidità; la crisi finale dovuta alla rottura delle fibre in trazione. Nel caso dei provini CP abbiamo solo due fasi, la prima relativa alla rottura delle lamine in compressione fino a circa la metà del momento a rottura dei provini non impattati seguita dalla fase di rottura di schianto delle fibre a trazione.

Conclusioni analoghe possono essere suggerite dall'ispezione visiva dei campioni portati a rottura. In figura 7 si possono osservare le lamine esterna ed interna di un provino CP dopo l'impatto e la torsione. Sono visibili nettamente sia la rottura di fibre a trazione sia l'ingobbamento delle fibre dovuto alla compressione. La linea di frattura passa per il punto di impatto, che è quello al di sotto del quale si localizza la delaminazione.

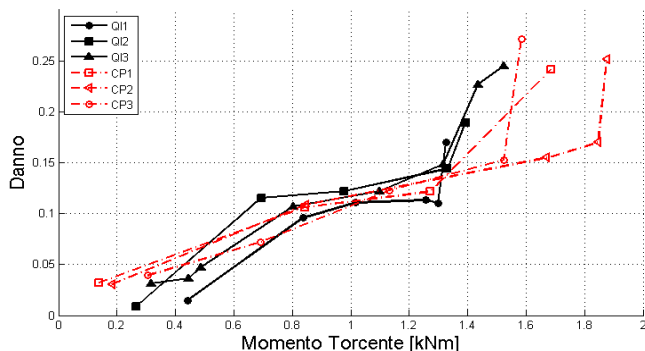


Figura 6. Evoluzione del danno in provini QI e CP

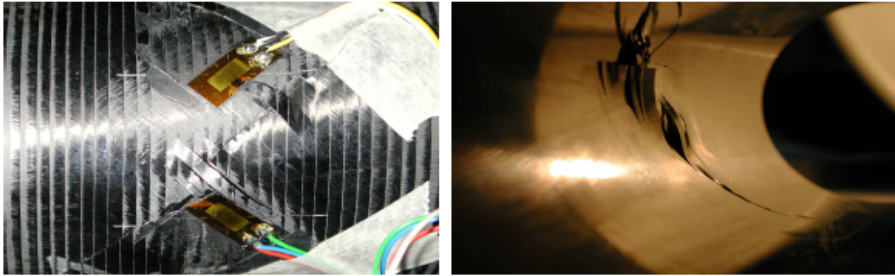


Figura 7. Rottura a torsione dopo l'impatto in campioni CP

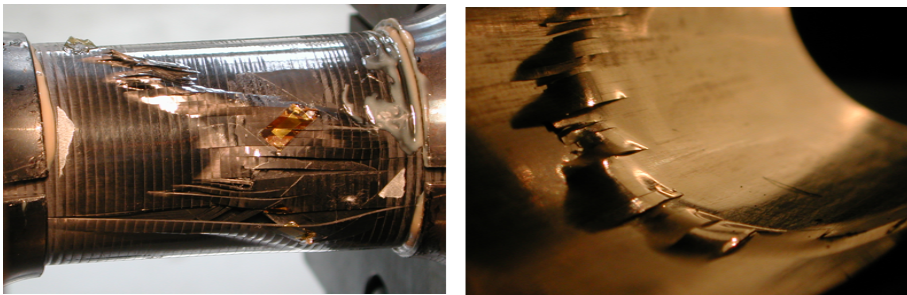


Figura 8. Rottura a torsione dopo l'impatto in campioni QI

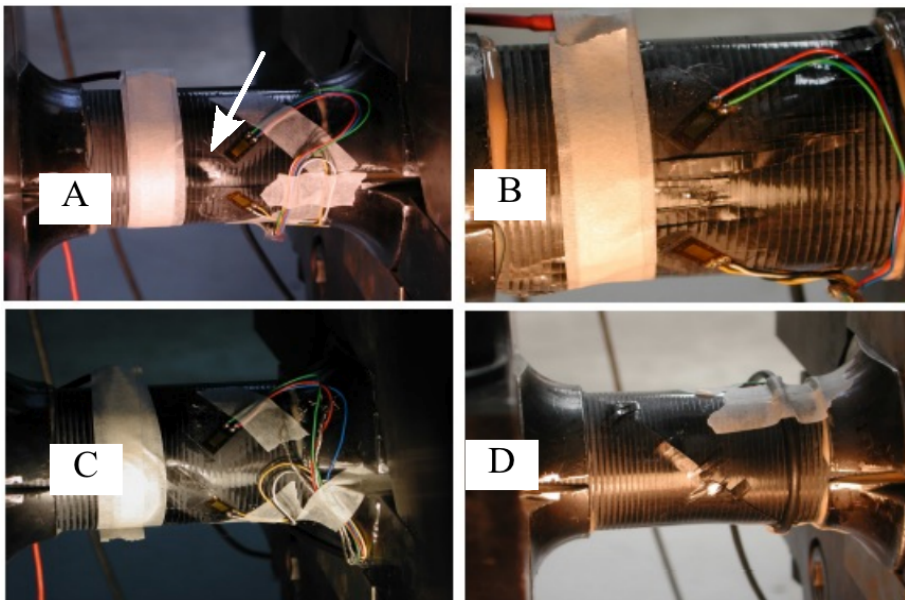


Figura 9. Impatto su campioni precaricati

In figura 8 sono mostrate le fotografie analoghe per il caso dei provini QI. La rottura è più complessa perché la frattura finale è mascherata dalla lamina esterna a 0°, fortemente delaminata e con varie rotture di matrice che si sviluppano nella direzione delle fibre.

Anche la lamina interna è caratterizzata da esteso splitting in corrispondenza della linea di frattura della lamina a 45° sottostante, soggetta a trazione.

Gli effetti dell'impatto su provini precaricati sono mostrati in figura 9. In questo caso abbiamo a disposizione solo l'immagine dell'esterno del tubo poiché dopo l'urto i provini sono stati portati immediatamente a rottura a torsione senza togliere il precarico e questo ha impedito di documentare il danno sulla faccia interna.

I tubi QI precaricati al 65% della RRT (figura 9A) hanno mostrato una marcata impronta, nettamente visibile, ed il distacco di parte della lamina esterna a 0° fratturata con splitting evidente.

Nel caso del precarico al 130% della RRT i campioni QI (figura 9B) hanno presentato una marcata rottura delle lamine interne che ha portato al distacco della lamina esterna in più punti (pur senza la rottura delle fibre poste a 0°).

Nei provini CP al livello più basso di precarico (figura 9C) si è prodotto un avvallamento centrato nel punto di impatto e orientato a 45° che nei tubi precaricati al livello più alto (figura 9D) è diventata una vera e propria frattura elicoidale. Questa frattura ha interessato l'intera lunghezza del tubo fino agli afferraggi e, come si è visto applicando una rotazione opposta a quella iniziale che provocava il precarico, ha attraversato tutto lo spessore.

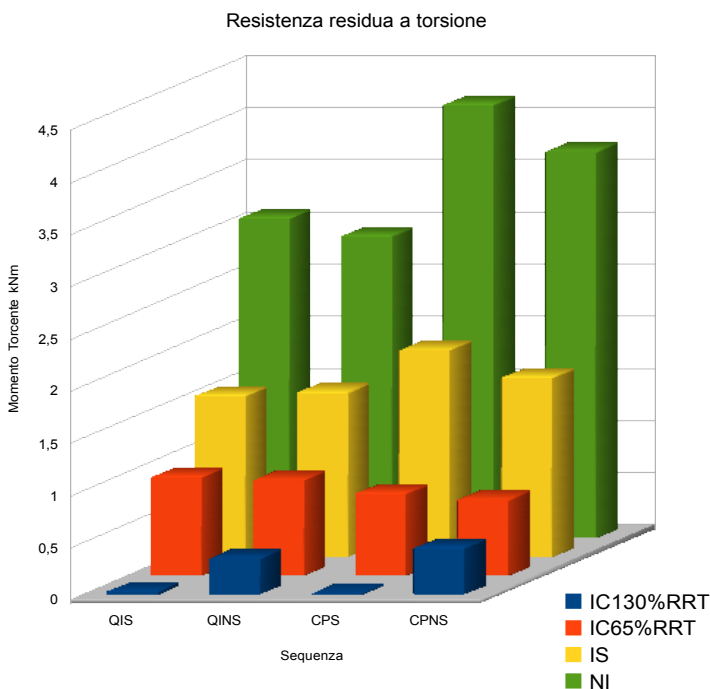


Figura 10. Grafico riassuntivo

La sperimentazione non ha dato motivi per pensare che vi siano differenze tra gli effetti di un'applicazione sequenziale o simultanea dell'impatto e della torsione. Al livello di precarico più basso in particolare, prima dell'impatto il provino risulta non danneggiato e l'effetto del precarico è solo quello di fornire l'energia necessaria allo sviluppo del danno provocato dall'impatto nei termini descritti nel caso della prove di resistenza residua dopo impatto. Al livello più elevato di precarico esistono già dei danni antecedenti all'impatto, ma sono localizzati nelle matrici e nelle zone adiacenti agli afferraggi, quindi lontani dal punto di impatto. Anche in questo caso l'energia immagazzinata provoca l'evoluzione del danneggiamento che può arrivare alla rottura visto che il momento applicato è superiore alla RRT.

In figura 10 è riportato il sommario delle momenti torcenti a rottura nelle varie configurazioni. Si osserva che non vi sono differenze sostanziali tra le laminazioni simmetriche e quelle non simmetriche e tra i provini impattati sotto precarico scompare anche la differenza tra QI e CP.

L'impatto, per quanto di energia notevolmente bassa, ed in grado di lasciare impronte appena visibili [8, 9], riduce la resistenza a torsione di oltre la metà.

Questo è coerente con quanto esposto in precedenza, infatti se eliminiamo tutte le lamine caricate in compressione, che si rompono prima di quelle in trazione a causa della delaminazione, otteniamo una riduzione di resistenza residua di quell'ordine di grandezza.

Ulteriori riduzioni si hanno nel caso dei provini impattati sotto carico, che nel caso del precarico più elevato portano praticamente a zero la resistenza residua.

4. CONCLUSIONE

Il comportamento a torsione di strutture tubolari in composito CFRP con differenti sequenze di laminazione, con e senza danno da impatto, è stato investigato sperimentalmente e si sono raggiunte le conclusioni seguenti.

La simmetria della sequenza di laminazione non influisce in maniera significativa sul comportamento a torsione e nel caso dei provini impattati è stata poca la differenza anche tra le sequenze QI e CP.

L'impatto, provocando una delaminazione, riduce il carico critico a compressione delle lamine e quindi la resistenza torsionale residua di almeno la metà. Analogamente un precarico torsionale durante l'impatto riduce drasticamente la resistenza residua anche se è inferiore al momento torcente che provocherebbe il primo danno di matrice. Un precarico superiore alla RRT porta alla rottura completa dei campioni durante l'impatto.

Appositi accorgimenti, quali il rivestimento con strati di materiale adatto ad assorbire l'energia in caso di impatti accidentali, quindi sono necessari per poter sfruttare le prestazioni superiori dei CFRP nelle applicazioni che prevedono forti sollecitazioni torsionali in esercizio.

Ringraziamenti

Gli autori ringraziano l'Ing. Dario Comand e la ditta REGLASS H.T. S.r.l di Minerbio (BO) per l'insostituibile supporto offerto nel corso dell'attività di ricerca e l'Ing. Riccardo Panciroli per la collaborazione nelle prove sperimentali.

BIBLIOGRAFIA

- [1] Cesari, F., Dal Re, V., Minak, G., Zucchelli, A., 2006. “Damage and residual strength of laminated carbon–epoxy composite circular plates loaded at the centre”, *Composites Part A: Applied Science and Manufacturing*, **38** (4), pp. 1163-1173.
- [2] Minak, G., Morelli, P., Zucchelli, A., 2008. “Fatigue residual strength of laminated graphite-epoxy composite circular plates damaged by transversal loads”, *Composite Science and Technology*, In Press 10.1016/j.compscitech.2008.05.025.
- [3] Ballere, L. , Viot, P., Lataillade, J.L., Guillaumat, L., Cloutet, S., 2008. “Damage tolerance of impacted curved panels”, *International Journal of Impact Engineering*, In Press 10.1016 /j.ijimpeng. 2008 .03.004.
- [4] Gning, P.B. Tarfaoui, M.; Collombet, F.; Davies, P., 2005. “Prediction of damage in composite cylinders after impact”, *Journal of Composite Materials*, **39** (10), pp. 917-28.
- [5] Bisagni, C, Cordisco, P., 2003. “An experimental investigation into the buckling and postbuckling of CFRP shells under axial and torsion loading”, *Composite Structures*, **60** (4), pp. 391–402.
- [6] Minak, G., Ghelli, D., Zucchelli, A., 2008. “Torsion Testing of CFRP laminate tubes”, Proceedings of the 5th ESIS TC4 International Conference on Fracture of Polymers, Composites and Adhesives, Les Diablerets.
- [7] Minak, G., Zucchelli, A., 2008. *Composite Materials Research Progress*, Novapublishers, New York, Capitolo.5:“Damage Evaluation and Residual Strength Prediction of CFRP Laminates by means of Acoustic Emission Techniques”, pp. 165-207.
- [8] Minak, G., Ghelli, D., Zucchelli, A., 2008. “Impatto a bassa velocità su strutture tubolari in materiale composito”, *ATTI XXXVII Convegno Nazionale AIAS*, ROMA disponibile sul sito http://www.pcm.unifi.it/AIAS2008/index_AIAS08.html.
- [9] Minak, G., Ghelli, D., Panciroli, R., Zucchelli, A., 2009. “Composite tubes behaviour under low velocity impact”, *Proceedings of MECSOL 2009*, Rio de Janeiro – Brazil.

THREE EXAMPLES OF INADEQUACY OF THE WILLIAMS ASYMPTOTIC METHOD IN PREDICTING THE SINGULARITY STRENGTH IN ELASTIC PROBLEMS

Antonio Strozzi

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: astrozzi@unimore.it*

Andrea Baldini

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: abaldini@unimore.it*

Matteo Giacomini

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: matteo.giacomini@unimore.it*

Roberto Rosi

*Department of Mechanical and Civil Engineering,
University of Modena and Reggio Emilia, Italy
E-mail: rrosi@unimore.it*

Abstract. *Three examples are presented for which the Williams asymptotic method appears to be inadequate in forecasting the strength of the singularity encountered in elastic problems. The first case addresses a Kirchhoff plate resting on an irregular linear support; the second example revisits a classical plane elasticity contact problem; the third case considers an infinite sheet with a square hole. The reasons of the inadequacy of the Williams asymptotic method in treating the above situations are discussed.*

Keywords: *Williams asymptotic approach, Kirchhoff plate, Contact problem, Reaction forces, Square hole, Singularity strength.*

1. INTRODUCTION

The Williams asymptotic method, e.g. [1–7], is an effective tool for forecasting the singularity strength of the stress field in elastic problems, e.g. Kirchhoff plates resting on irregular linear supports and plates stretched in their plane and exhibiting re-entrant corners.

One of the prerogatives of the Williams method is its ability to predict the singularity strength of the stress field without requiring the complete solution of the problem under

scrutiny It is however noted that the Williams method rests on a homogeneous solution of the biharmonic equation of elasticity, and that it does not account for particular solutions. In other words, the singularity strength predicted by the Williams method is independent of the particular loading or constraining. Since the solution adopted in the Williams approach is incomplete, it may be suspected that for particular problems the Williams approach might supply incorrect predictions on the singularity strength. In other words, the Williams approach may become inadequate in peculiar circumstances.

This study presents three particular problems in which the Williams method is inadequate in forecasting the elastic singularity strength. The first problem deals with an infinite Kirchhoff plate resting on an irregular linear support; the second problem refers to a plane elasticity contact problem; the third example addresses an infinite sheet with a square hole, loaded in its plane. The three examples of inadequacy of the Williams approach rest on a singular solution of the stress field under the effects of a certain system of forces, where the singularity strength agrees with the Williams predictions. A second set of forces is then considered which, when acting simultaneously to the first set, osculates the singularities originated by the first set, thus producing an example which disagrees with the Williams forecasts. The present authors are unaware of specific studies evidencing situations of inadequacy of the Williams approach.

2. THE WILLIAMS ASYMPTOTIC METHOD

The Williams asymptotic method rests on the following homogeneous form of the classical biharmonic equation:

$$\Delta\Delta\Psi = 0 \tag{1}$$

where Ψ represents the plate deflection in problems of bent plates, as [1], or the Airy stress function in problems of plane elasticity, as in [2]. Williams adopts the following polar form of the solution to (1):

$$\Psi(r, \theta) = r^{\lambda+1} [b_1 \sin(\lambda + 1)\theta + b_2 \cos(\lambda + 1)\theta + b_3 \sin(\lambda - 1)\theta + b_4 \cos(\lambda - 1)\theta] \tag{2}$$

where the origin of the polar coordinates is centered on the transition point for the boundary conditions. In practice, the origin coincides with a support extremity in plates resting on linear supports, and with the vertex of the corner in plane elastic problems with sharp notches. The imposition to (2) of suitable homogeneous boundary conditions results in a set of four homogeneous algebraic equations, which are linear in the four coefficients b_i of (2). The existence of non trivial solutions requires that the determinant of the coefficients b_i vanish, a condition that allows the eigenvalues of the exponent λ of (2) ($0 < Re(\lambda) < 1$) to be determined and, consequently, the singularity strength to be predicted, see [1].

Infinite Kirchhoff plate resting on a linear discontinuous support

A review of the singularity strength of the contact reaction in Kirchhoff plates resting on irregular linear supports has recently been carried out in [8]. The contact reaction may be expressed in terms of a Kirchhoff equivalent shear force, condensing the twisting moment effects on the plate deflection into the shear force. According to the Williams approach, this

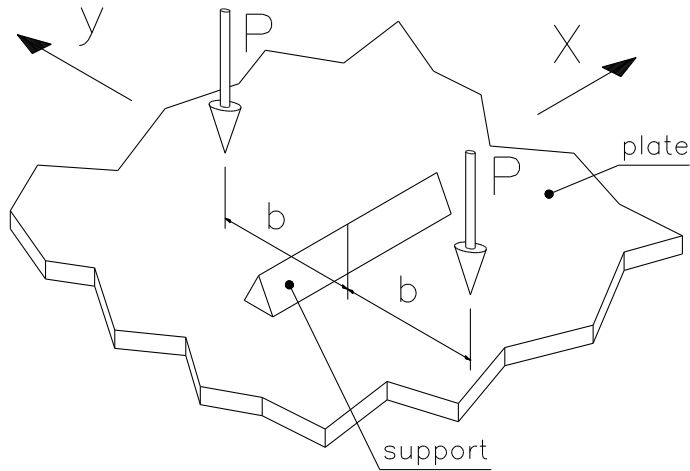


Figure 1: Infinite Kirchhoff plate resting on a linear segment and loaded by a single pair of concentrated transverse forces.

shear force exhibits a non integrable singularity strength of the kind $1/(x\sqrt{x})$. If, conversely, the contact reaction is described in terms of a non traditional "Kirchhoff equivalent twisting moment", see [8] for further details, condensing the shear force effects into the twisting moment, the endpoint singularity strength of the twisting moment is integrable and of the kind $1/\sqrt{x}$.

In [8] an exact solution has recently been obtained for an infinite Kirchhoff plate resting on a linear segment of length 2 and loaded by a single pair of concentrated transverse forces P , symmetrically applied at both sides of the support, with an offset denoted by b ($b > 0$). Figure 1 clarifies the plate loading, and specifies the orientation of the Cartesian axes x and y .

The contact reaction expressed in terms of a skew-symmetric equivalent twisting moment m is

$$m(x) = \frac{2P}{\pi} \left[\operatorname{atan} \left(\frac{x\sqrt{1+b^2}}{b\sqrt{1-x^2}} \right) + \frac{xb\sqrt{1+b^2}}{\sqrt{1-x^2}(b^2+x^2)} - \frac{x \ln(b + \sqrt{1+b^2})}{\sqrt{1-x^2}} \right] \quad (3)$$

The distributed twisting moment m is accompanied by two endpoint concentrated forces, each of intensity P , which equilibrate the imposed loads. From expression (3) it appears that, for any positive offset b , the distributed twisting moment m exhibits an endpoint singularity of the kind $1/\sqrt{x}$, as correctly predicted by the Williams method.

The contact reaction (3) may be converted (through integration by parts, as in [8]) into a distributed shear force accompanied by endpoint, infinite, equilibrating, concentrated forces; for the particular loading detailed above, such distributed shear force possesses an endpoint singularity strength of the kind $1/(x\sqrt{x})$, as correctly predicted by [1].

In the following, a variant of the loading described in Fig. 1 is discussed, aimed at evidencing the limits of the Williams approach in peculiar situations. This new loading is formed by two pairs of forces; the first pair is constituted by two equal forces P_1 straddling the support midpoint and possessing an offset b_1 , whereas the second pair is formed by two equal forces P_2 straddling the support midpoint, and exhibiting an offset b_2 .

The global contact reaction, m_g , is obtained by superposing the two corresponding solutions (3)

$$m_g(x) = \frac{2P_1}{\pi} \left[\operatorname{atan} \left(\frac{x\sqrt{1+b_1^2}}{b_1\sqrt{1-x^2}} \right) + \frac{xb_1\sqrt{1+b_1^2}}{\sqrt{1-x^2}(x^2+b_1^2)} - \frac{x \ln(b_1 + \sqrt{1+b_1^2})}{\sqrt{1-x^2}} \right] + \frac{2P_2}{\pi} \left[\operatorname{atan} \left(\frac{x\sqrt{1+b_2^2}}{b_2\sqrt{1-x^2}} \right) + \frac{xb_2\sqrt{1+b_2^2}}{\sqrt{1-x^2}(x^2+b_2^2)} - \frac{x \ln(b_2 + \sqrt{1+b_2^2})}{\sqrt{1-x^2}} \right] \quad (4)$$

It appears that the distributed twisting moment m_g does not exhibit any endpoint singularity if

$$\frac{P_1}{P_2} = - \frac{\frac{b_2}{\sqrt{1+b_2^2}} - \ln(b_2 + \sqrt{1+b_2^2})}{\frac{b_1}{\sqrt{1+b_1^2}} - \ln(b_1 + \sqrt{1+b_1^2})} \quad (5)$$

If condition (5) is verified, the endpoint values of the skew-symmetric reaction m_g are finite and equal to $\pm(P_1 + P_2)$. For positive b_1 and b_2 offsets, the ratio P_1/P_2 remains finite and negative, thus indicating that the orientation of the two pairs of forces must be opposite. In addition, $P_1/P_2 = -1$ only when $b_1 = b_2$, which represents a trivial undeformed situation.

It may be concluded that, for the particular loading constituted by two pairs of forces straddling the midpoint of the linear support with offsets whose relative intensity fulfil condition (5), the contact reaction expressed in terms of distributed twisting moment does not possess an endpoint singularity strength of the kind $1/\sqrt{x}$, as forecast by Williams, but it remains finite at the contact endpoints. In other words, the particular linear combination (4) of two solutions (3) that fulfils condition (5), results in a reduction of the singularity strength of the contact reaction with respect to a single solution (3).

If the contact reaction (4) expressed in terms of distributed twisting moment is changed into a distributed shear force accompanied by endpoint concentrated forces, for the particular loading detailed above such distributed shear force would possess an endpoint singularity strength of the kind $1/\sqrt{x}$, in contrast with [1].

A simple numerical example is carried out in the following. It is convenient to impose different numerical values to the offsets b_1 and b_2 , and to determine from expression (5) the corresponding value of the ratio P_1/P_2 . For instance, if $b_1 = 10$ and $b_2 = 11$, from expression (5) the ratio $P_1/P_2 \approx -1.047$ is derived. Consequently, for a loading formed by a) two concentrated forces P defined by an offset with respect to the support midpoint of 10, whose intensity is 1.047, and b) two concentrated forces P defined by an offset of 11, whose intensity is 1, and whose orientation is opposite to that of a), the contact reaction expressed in terms of distributed shear force exhibits integrable endpoint singularities, where this result contrasts with the [1] predictions.

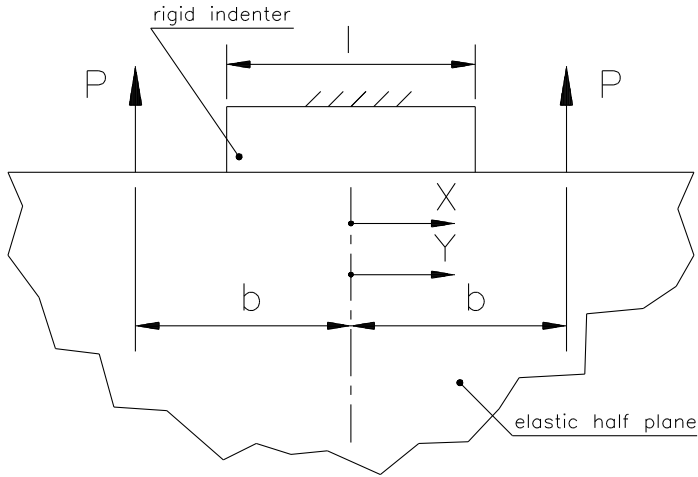


Figure 2: Deformable half plane compressed against a rigid rectangular indenter by a pair of transverse concentrated forces.

Yang [9] considers the contact between a rectangular plate and a central linear support; the contact reaction is expressed in terms of a pure linear shear force exhibiting integrable endpoint singularities. In [10] doubts are expressed on the correctness of the [9] results, "since the singularity is not in agreement with" the body of the literature studies, supporting stronger endpoint singularities. The results of [11] and of [8] confirm that the study of [9] is generally incorrect. However, the Kirchhoff plate example developed in this section indicates that in some peculiar circumstances the contact reaction along a linear support may be formed by a pure contact reaction force with integrable endpoint singularities. Consequently, in the [9] configuration it cannot be excluded *a priori* that for particular proportions of the rectangular plate and of the linear support the reaction force might exhibit integrable endpoint singularities.

The inadequacy of the Williams approach in forecasting the singularity strength of the contact reaction in Kirchhoff plates resting on irregular linear supports, may be confidently attributed to the Williams asymptotic approach neglecting the effects of the particular solution (connected to the plate loading). In the example developed in this study, a peculiar plate loading produces a diminution of the singularity strength of the contact reaction.

A contact problem in plane elasticity

The classical frictionless contact problem between a deformable half plane and a rigid rectangular indenter is revisited in the light of the possible existence of exceptions to the Williams predictions in terms of endpoints singularities. The half plate is compressed against the rigid rectangular indenter of length $l=2$ by two equal transverse concentrated forces P , applied to the half plane border with an offset of b ($b > 1$) with respect to the contact midpoint, Fig. 2.

By employing the following classical Green function, e.g. [12, p. 17],

$$w(x) = -\frac{2(1-\nu^2)}{\pi E} P \ln x \quad (6)$$

the integral equation describing this contact problem is

$$\int_{-1}^{+1} p(y) \ln|x-y| dy = P(b^2 - x^2) \quad (7)$$

where the coordinates x and y (dummy variable) are shown in Fig. 2. By differentiating the integral equation (7) once with respect to x , the canonical Cauchy integral equation is obtained

$$\int_{-1}^{+1} \frac{p(y)}{x-y} dy = -\frac{2Px}{b^2 - x^2} \quad (8)$$

By accounting for the following integral, which is a variant of result (33) of [8]

$$\int_{-1}^{+1} \frac{y\sqrt{1-y^2}}{(b^2-y^2)(x-y)} dy = -\pi \left(1 - \frac{b\sqrt{b^2-1}}{b^2-x^2}\right) \quad b > 1 \quad (9)$$

the solution of the integral equation (7) is, e.g. [13, p. 178]

$$p(x) = \frac{C}{\sqrt{1-x^2}} - \frac{1}{\pi} \frac{2P}{\sqrt{1-x^2}} \left(1 - \frac{b\sqrt{b^2-1}}{b^2-x^2}\right) \quad (10)$$

By imposing the vertical equilibrium condition

$$\int_{-1}^{+1} p(x) dx = 2P \quad (11)$$

the constant C of (10) is found to be

$$C = \frac{2P}{\pi} \quad (12)$$

so that the contact pressure p becomes

$$p(x) = \frac{2Pb\sqrt{b^2-1}}{\pi\sqrt{1-x^2}(b^2-x^2)} \quad (13)$$

The contact pressure (13) exhibits the typical endpoint square root singularity, as correctly predicted by the Williams method. If, however, two systems of forces are simultaneously applied, where the first system is formed by two forces P_1 defined by an offset b_1 , whereas the second system is constituted by two forces P_2 possessing an offset b_2 , the combined contact pressure is the sum of the two contributions (a bilateral contact reaction is assumed), and this pressure does not become infinite at the contact endpoints if the following combination between force intensity and offset is satisfied

$$\frac{P_1}{P_2} = -\frac{b_2\sqrt{b_1^2-1}}{b_1\sqrt{b_2^2-1}} \quad (14)$$

For the particular combination expressed by equation (14), the contact pressure is

$$p(x) = \frac{2P_1 b_1 (b_1^2 - b_2^2) \sqrt{1 - x^2}}{\pi \sqrt{b_1^2 - 1} (b_1^2 - x^2) (b_2^2 - x^2)} \quad (15)$$

The contact pressure (15) is null at the contact ends, a result which contrasts with the Williams asymptotic predictions. Consequently, when the above specific system of forces is employed in this plane elastic contact problem, the Williams approach appears to be inadequate. The reasons of this inadequacy have already been discussed in section 2.1.

An infinite sheet with a square hole

In this section the plane elasticity problem is considered of an infinite, square holed sheet loaded by a remote uniform tensile stress σ_1 as shown in Fig. 3 (a). According to [1, 2], the stress field in the vicinity of a re-entrant corner of the square hole may be expressed in terms of an eigenfunction expansion, [4, p. 144]; only the first term of the eigenfunction expansion is singular, see [3, p. 186], [4, p. 148]; the singularity strength of the stresses is of the kind $1/r^{(\lambda-1)}$, where λ ($0 < Re(\lambda) < 1$) is the solution of the characteristic equation for the symmetric problem

$$\lambda \sin(2\alpha) + \sin(2\lambda\alpha) = 0 \quad (16)$$

where the angle 2α describes the angular extent of the material surrounding the corner of the square hole, $2\alpha = 3\pi/2$, so that $\lambda \approx 0.5445$.

If the sheet is loaded by a remote uniform tensile stress σ_1 as shown in Fig. 3 (a), the two re-entrant corners labelled *A* tends to open, whereas if the sheet is stretched by a remote uniform tensile stress σ_2 as shown in Fig. 3 (b), the same corners *A* tends to close. The strength of the stress singular term is the same for the two above loadings of Fig. 3 (a) and (b), whereas its weight is generally different. Similar to the two examples developed in the previous sections 2.1 and 2.2, it may be surmised that a proportion exists between the intensity of the two above loadings of Fig. 3 (a) and (b) for which the stress singularities activated by the two loadings osculate each other at the two corners labelled *A*, whereas the stresses remain infinite at the two remaining corners. Straightforward finite element calculations indicate that this critical ratio σ_2/σ_1 is about 4.6. (For slightly different load proportions, all the edge stresses become again infinite, so that it would be virtually impossible to employ this stress reduction in engineering applications.)

Although the lack of a closed form solution for the square hole problem precludes the exact determination of the critical loading ratio σ_2/σ_1 , the information gathered is sufficient to conclude that a very peculiar loading exists for which the stresses remain finite at the two re-entrant sharp corners lettered *A*. This result disagrees with the Williams approach, which predicts infinite stresses in re-entrant sharp corners of a plate loaded in its plane, independent of the loading imposed, e.g. [2]. Consequently, in this example too the Williams approach appears to be inadequate.

The above square hole problem has been spurred by a similar elastic problem analysed in [14, p. 305]. The stress concentrations have been examined photoelastically in an infinite sheet exhibiting a square hole with sharp corners, and subjected to remote uniform shear at 45° to the sides of the square, Fig. 3 (c); the tips of the corners have been photoelastically found by Heywood to be low-stressed regions.

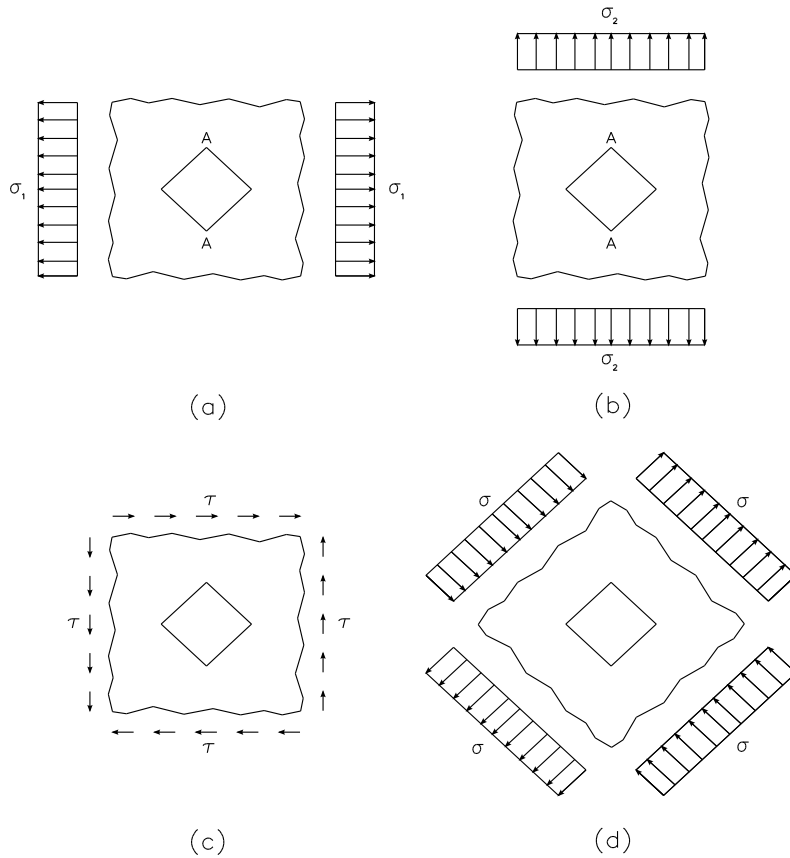


Figure 3: Infinite square holed sheet loaded by a combination of uniform remote stresses

It is noted that the remote pure shear loading of Fig. 3 (c) may be interpreted as a remote biaxial loading in which the intensity of the compression forces equals that of the tensile loads, Fig. 3 (d). Contrary to the situations analysed in sections 2.1 and 2.2, a degree of freedom expressing a variable proportion between two loadings is not available in the case of Fig 3 (d), so that it is unlikely that this loading produces stress singularity oscillation and, consequently, finite stresses at the corner tips. The present authors in fact believe that the Heywood results are spurious in the case of sharp corners. The photoelastically low stresses noted by Heywood may be attributed to the circumstance that the corners of the photoelastic model are inevitably rounded, so that the photoelastic stresses are null at the midpoints of the arcs defining the rounded tips, e.g. [15, formula 42] and related comments. (The present authors have replicated the photoelastic study of Heywood, and they have found the same fringe pattern of [14, Fig. 8.24].)

3. CONCLUSIONS

Three problems have been examined that, under the effects of a particular loading, exhibit endpoint stress singularities whose strength disagrees with the Williams forecasts. The first problem addresses the contact between an infinite Kirchhoff plate and a linear discontinuous support, for which an exact solution in terms of contact reaction was available. Although for general loadings the endpoint singularity strength of the contact reaction agrees with the Williams prediction, a peculiar loading has been detected for which the singularity strength is lower than the Williams forecast. The second problem deals with a classical contact problem in plane elasticity. For a particular combined loading, the endpoint singularity strength disagrees with the Williams predictions. The third problem addresses the plane elastic problem of an infinite sheet with a square hole. For a particular loading, the stresses do not become infinite at two of the four sharp corners of the square hole, whereas the Williams method would predict infinite stresses at all four corners.

It is noted that, as soon as the peculiar loading that produces stress singularities disagreeing with the Williams predictions is slightly perturbed, the strength of the stress singularities turns to agree with the Williams forecasts. Consequently, it is practically impossible to employ the above results to lower the stress concentration level in engineering applications.

It may be concluded that in some very peculiar situations the Williams asymptotic method may be inadequate in predicting the strength of the stress singularity. The reasons for this inadequacy have been attributed to the Williams approach neglecting the effects of a particular solution dealing with the imposition of loading and/or constraining.

REFERENCES

- [1] Williams, M. "Surface stress singularities resulting from various boundary conditions in angular corners of plates under bending". In US National Congress of Applied Mechanics, pp. 325–329.
- [2] Williams, M., 1952. "Stress singularities resulting from various boundary conditions in angular corners of plates in extension". *J. Appl. Mech.*, **19**(4), pp. 526–528.
- [3] Soutas-Little, R., 1999. *Elasticity*. Courier Dover Publications.
- [4] Barber, J., 2002. *Elasticity*. Kluwer Academic Publishers.
- [5] Szilard, R., 2004. *Theories and Applications of Plate Analysis: Classical Numerical and Engineering Methods*. Wiley.
- [6] Sinclair, G., 2004. "Stress singularities in classical elasticityII: Asymptotic identification". *Applied Mechanics Reviews*, **57**, p. 385.
- [7] Lee, D., and Barber, J., 2006. "An Automated Procedure for Determining Asymptotic Elastic Stress Fields at Singular Points". *The Journal of Strain Analysis for Engineering Design*, **41**(4), pp. 287–295.
- [8] Strozzi, A., and Monegato, G., 2008. "On the incompatibility between the equivalent shear force concept and the integral formulation of contact problems between Kirchhoff plates and irregular linear supports". *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **222**(7), pp. 1149–1163.
- [9] Yang, W., 1968. "On an integral equation solution for a plane with internal support". *The Quarterly Journal of Mechanics and Applied Mathematics*, **21**(4), pp. 503–515.
- [10] Stahl, B., and Keer, L., 1972. "Vibration and buckling of a rectangular plate with an internal support". *The Quarterly Journal of Mechanics and Applied Mathematics*, **25**(4), pp. 467–478.

- [11] Sompornjaroensuk, Y., and Kiattikomol, K., 2008. “Exact analytical solutions for bending of rectangular plates with a partial internal line support”. *J. Eng Math.*
- [12] Johnson, K., 1987. *Contact Mechanics*. Cambridge University Press.
- [13] Tricomi, F., 1985. *Integral Equations*. Courier Dover Publications.
- [14] Heywood, R., 1952. *Designing by Photoelasticity*. Chapman & Hall.
- [15] Filippi, S., Lazzarin, P., and Tovo, R., 2002. “Developments of some explicit formulas useful to describe elastic stress fields ahead of notches in plates”. *International Journal of Solids and Structures*, **39**(17), pp. 4543–4565.

DESIGN DRIVERS PER LO SVILUPPO DI NUOVO PRODOTTO

Gianluca Medri
Department of Industrial Engineering,
University of Parma, Italy
E-mail: gianluca.medri@unipr.it

Abstract. *This paper introduces the concept of design drivers, as fundamental instruments and guidelines to lead the design process of a new product up to useful and reliable results. An example of design driver identification is presented, with reference to the problem of the arrangement of user-product interfaces.*

Sommario. *Questa memoria introduce il concetto di design drivers, strumenti fondamentali e linee guida per condurre il processo di progettazione di un nuovo prodotto a risultati utili ed affidabili. Si presenta un esempio di identificazione di design drivers in riferimento al problema dell'organizzazione delle interfacce utente-manufatto.*

Keywords: *Design drivers, new product development, design process.*

1. IL CONCETTO DI DESIGN DRIVER

L'indirizzo generale per la progettazione di un nuovo prodotto è basato sull'ottenimento di miglioramenti funzionali, tecnico-economici ed estetici (per quanto da quest'ultimo punto di vista si possa fare riferimento alla funzionalità intesa in senso lato) rispetto all'esistente. Conseguentemente, lo sviluppo di nuovo prodotto deve essere basato su un'adeguata e ben definita strategia d'impresa, un'organizzazione aziendale efficiente e deve riferirsi a indicazioni che incanalino la progettualità verso soluzioni funzionali, sicure, tecnologicamente efficienti e ad elevata probabilità di successo sul mercato.

Per alcune tipologie di queste indicazioni si propone la definizione di **design drivers**, per distinguerle concettualmente dalla generica o specifica normativa tecnico-legale e dalle più specifiche (nel senso di più indirizzate a particolari tipologie di manufatti e più approfondite nei contenuti) linee-guida (**guidelines**). Si presentano, di seguito, sintetiche considerazioni sulla progettazione e sulla definizione di *design drivers*.

Nella pratica industriale, la spinta creativa dei progettisti (*designers* e tecnici progettisti costituenti il **Product Development Team** - PDT) viene guidata, seguendo procedure sistematiche o istintuali (in questo secondo caso, vezzo proprio del cosiddetto *Design*), all'innovazione di forma, di materiali e di soluzione tecnica (*concept*) con il pensiero costantemente orientato anche al rispetto delle compatibilità economiche ed ambientali. Questo è l'inquadramento del processo progettuale relativo agli obiettivi generali progettuali (metaobiettivi: funzionalità, sicurezza, economicità), ai quali si aggiungono normalmente gli obiettivi specifici che permettono di inquadrare il prodotto (target di prodotto, di mercato etc.).

Si accenna brevemente al fatto che la visione strategica dell'impresa è tradizionalmente

incentrata sul presente e, al massimo, su un futuro prossimo a questo meccanicamente collegato. In questa ottica il passato viene usato come *data-base* e come riferimento per la valutazione dei *trend* critici; quando si effettuano analisi predittive economiche a lungo termine queste sono generalmente basate sulla sostanziale stabilità delle condizioni al contorno, o al massimo su variazioni quantitativamente ridotte di queste ultime con esclusione di significative variazioni qualitative. Questo modo di procedere rende solida, veloce e relativamente semplice l'attività decisionale e progettuale, perché si fa riferimento solo all'esistente (unico ed oggettivo), e risulta più rassicurante di una visione basata su predizioni. D'altra parte è evidente la sostanziale miopia conservatrice di un siffatto atteggiamento: **si progetta per il futuro un prodotto pensato per il presente** che sarà già concettualmente sorpassato quando entrerà in produzione (e non solo per il *lead time* o per il naturale progredire delle conoscenze tecnologiche, ma anche per la possibile variazione delle richieste di prodotti e servizi). Lo scenario futuro deve, invece, essere investigato in modo (obbligatoriamente) probabilistico e qualitativamente connotato, prefigurando evoluzioni possibili (*future studies*) sui fronti economico, politico-sociale, culturale, tecnologico ed antropologico. Questa procedura fornisce informazioni più adeguate per una **strategia di progettazione in prospettiva** (Barquero Perez, 2008) tale che il prodotto viene progettato per il futuro e contribuisce esso stesso a generare il futuro con la sua carica di innovazione, chiudendo il ciclo temporale metaprogettuale. Naturalmente i futuri possibili sono molti e diversi e l'attività decisionale di pianificazione progettuale risulta più complessa e responsabilizzata dall'aleatorietà dello scenario di mercato.

I *design drivers* quindi derivano

- dai metavincoli progettuali del settore merceologico di riferimento,
- dalle conoscenze tecnologico-progettuali del settore industriale specifico,
- dalle necessità del cliente/utente in riferimento allo scenario strategico prospettico sopra accennato,

e possono essere sensibilmente falsati dalla mancanza di questo scenario.

Essi costituiscono un'importante parte del *know-how* condiviso o proprietario delle imprese in ogni specifico settore merceologico e in ogni distretto industriale.

Per quanto detto, i *design drivers* si aggiungono agli obiettivi e ai classici vincoli progettuali, generali (Pahl e Beitz [1]: sicurezza, ergonomia, produttività, fabbricazione, trasporto, uso, manutenzione, costo; con l'aggiunta della sostenibilità) e specifici del prodotto, ai quali si aggiunge l'**estetica** nell'*Industrial Design*.

2. IDENTIFICAZIONE DI *DESIGN DRIVERS*

L'identificazione dei *design drivers* specifici per uno sviluppo di nuovo prodotto fa riferimento all'esperienza e alle conoscenze tecniche-industriali-funzionalistiche, alla normativa tecnica e legale e, come già detto, alla strategia d'impresa. E', ad esempio, un basilare *design driver* il principio dell'integrazione della sicurezza indicato nella Direttiva Macchine, che privilegia il prodotto intrinsecamente sicuro rispetto a quello che adotta sistemi di sicurezza indiretta. Si riporta di seguito, a titolo di esempio, il processo di riconoscimento di *design drivers* per manufatti a sensibile interazione operativa prodotto/utente.

L'evoluzione culturale e socio-economica ha (ri)portato ad una attenzione rilevante verso la sicurezza e la garanzia di funzionalità dei manufatti e da questo si può partire per indicazioni metaprogettuali che permettano di integrare a priori sicurezza-ergonomia-uso (e

anche, in molti casi, facilità di montaggio e manutenzione) nel prodotto. Riconoscendo che, nella realtà operativa, l'unità funzionale è costituita dal manufatto e dall'utente-fruitore-operatore (fig. 1), è compito dei progettisti ottimizzare questa unità nella progettazione che, quindi, deve toccare e coinvolgere anche l'utente-fruitore-operatore. Questo concetto è un fondamentale metavincolo progettuale, che contribuisce a generare efficienti *design drivers* per molte categorie di manufatti.

Incidentalmente, per la struttura artigianale della produzione di manufatti, nell'epoca pre-industriale il concetto del sistema manufatto-utente era un'ovvia conseguenza (oltre a far parte del bagaglio professionale standard di ogni artigiano) del contatto diretto (o molto meno mediato) tra cliente-utente e produttore-progettista.

Lo schema strutturale, molto semplice ed intuitivo ma contemporaneamente molto pregnante, di fig. 1 identifica il confine, verso l'esterno – qualunque cosa ciò significhi, del sistema “unità funzionale” con l'ovale a linea grossa e l'**interfaccia utente-manufatto** (tutta **interna** al sistema funzionale) con la linea sottile. In una visione funzionale brutalmente e rozzamente meccanicistica, ma molto didattica, l'unità funzionale si può, quindi, considerare costituita da una sotto-unità operativa, da una sotto-unità di potenza e da una sotto-unità di governo; queste che possono “appoggiarsi” sia sul manufatto sia sull'utente che diventano specifici *function carriers* dell'unità stessa.

A titolo esemplificativo, lo strumento *schiaccianoci* è costituito dalla sotto-unità operativa “schiaccianoci” (il simpatico strumento casalingo che ci permette di gustare il gheriglio delle noci senza rovinarci i denti e le mani!), dalla sotto-unità di governo “cervello dell'utente” e dalla sotto-unità di potenza “mani dell'utente” (sono possibili ovviamente altre identificazioni funzionali): in questo caso l'interfaccia utente-manufatto separa la sotto-unità operativa dalle altre due. In altri casi la sotto-unità di potenza è interna al manufatto etc. etc. ma, in ogni caso e per ogni prodotto, si ripete che

l'interfaccia utente-manufatto è interna al sistema funzionale e questo costituisce un fondamentale *design driver* della progettazione in generale

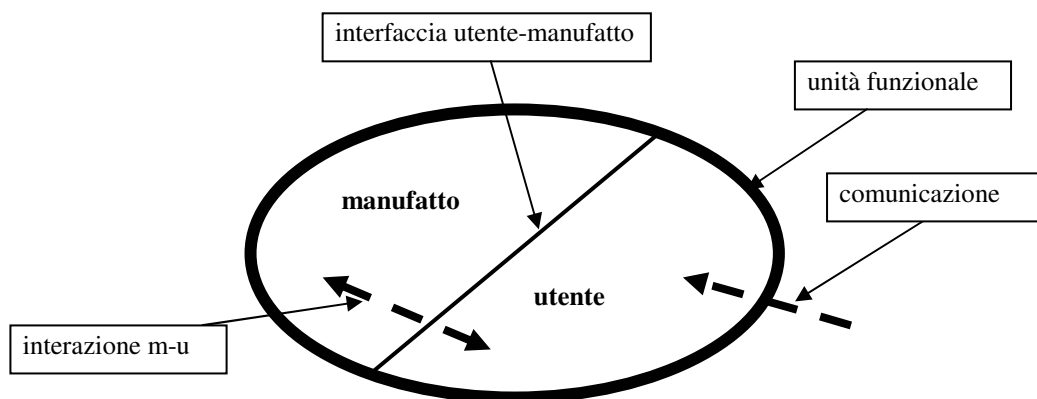


Figura 1. Il sistema funzionale utente-manufatto.

Il significato di questo *design driver* travalica la semplice classificazione tassonomica del sistema funzionale, perché prescrive l'integrazione funzionale dell'utente e del prodotto!

La linea di confine del sistema è attraversata da segnali (fisico-sensoriali, chimici o informativi) che rappresentano l'interazione con l'esterno. Se il sistema funzionale opera su oggetti o materiali (non necessariamente a contatto fisico nel senso usuale del termine) questi retroagiscono, ovvero da questi partono segnali informativi e l'insieme delle informazioni-retroazioni permette all'unità operativa di funzionare al meglio (se le informazioni e i segnali sono elaborati ed interpretati correttamente). La rappresentazione di fig. 1 è qualitativa e simbolica anche in riferimento alla vita di servizio del manufatto, in quanto l'interfaccia utente-manufatto (e il confine sistema-ambiente) può, in linea di principio, variare quantitativamente e qualitativamente (sia in senso topologico sia in senso fisico) nel tempo o in fasi diverse dell'utilizzo, naturalmente sempre in modo previsto e controllato in sede progettuale.

Nella progettazione il PDT può intervenire sul componente "utente" (sulle sue varie funzionalità: di governo, di potenza od operativa) attraverso due vie:

- la comunicazione che attraversa il confine esterno del sistema, può essere diretta e indiretta (quest'ultima mediata attraverso, o specifica per, l'oggetto passivo della funzione),
- l'interazione manufatto-utente (che passa attraverso l'interfaccia interna).

L'efficacia degli interventi progettuali in questo campo dipende dalla capacità di comunicazione e di manipolazione (nel senso positivo del termine) nei riguardi dell'utente che, in linea di principio, non è facilmente identificabile nella sua completezza di caratteristiche fisiche e intellettuali. Classici esempi (indicativi e non esaustivi) di interventi progettuali sull'utente sono

comunicazione diretta (temporalmente indipendente dallo stato operativo del manufatto)

- messaggi pubblicitari,
- la grafica della confezione,
- il Manuale di Uso e Manutenzione.

comunicazione indiretta (in esercizio)

- misure di grandezze fisiche dell'ambiente (modificabili dall'azione dell'unità funzionale),
- segnali dal manufatto verso l'utente (generati dall'interazione con l'ambiente o con l'oggetto passivo della funzione).

interazione manufatto-utente: inserimento sul prodotto di specifici

- organi di controllo e comandi,
- configurazioni strutturali,
- protezioni,
- informazioni
- segnali o avvertimenti scritti o pittografici,
- sistemi *mistake-proofing*,
- forme, odori, suoni e colori per ottenere attenzione o reazioni istintive.

Il PDT ha, ovviamente, la possibilità di intervenire più fisicamente e più compiutamente sul manufatto che sull'utente. Il manufatto, infatti, non deve essere convinto

o istruito ma viene costruito definendone caratteristiche, architettura e funzionamento! Tuttavia, si ribadisce che le scelte progettuali devono considerare e contemperare obiettivi e vincoli operando su entrambe le componenti del sistema funzionale, e questo può essere facilitato (da un certo punto di vista) dall'individuazione preliminare dei *design drivers* che caratterizzano il settore merceologico e la specifica funzione del manufatto da progettare. I concetti da applicare nell'*embodiment* strutturale sono mutuabili dalla Progettazione Modulare (vedi [2]), considerando l'utente come modulo intercambiabile e curando quindi in modo particolare la funzionalità di collegamento con gli altri moduli del sistema (i componenti del manufatto e gli oggetti passivi della funzionalità di questo).

Una delle più interessanti categorie di *design drivers* è collegata alla configurazione fisico-sensoriale dell'interfaccia utente-manufatto, quando questa sia stabilita (e non alterabile se non a prezzo di significative riduzioni di ergonomia, sicurezza e funzionalità d'uso, caratteristiche che permettono di identificare un manufatto come *user friendly*) da un *embodiment* ottimizzato dall'uso e dal tempo (e, si noti, che sia sempre giustificabile secondo vincoli e obiettivi progettuali).

Questa specifica analisi è particolarmente riferibile (ma non solo!) alla progettazione e sviluppo di nuovi prodotti destinati a funzioni di servizio **definite e non innovative**. Rientrano in questa categoria:

- manufatti migliorati o ottimizzati rispetto a quelli già esistenti, ma da questi derivati (progettazione adattativa secondo Pahl e Beitz [1]);
- manufatti di nuova concezione (forma o soluzione tecnica o materiali - etc.-innovativi) per funzioni note e già erogate da manufatti esistenti.

In questo caso, poiché l'alterazione dell'interfaccia comporta in generale un peggioramento dell'insieme delle caratteristiche volute e desiderabili del manufatto (*product design requirements* - **PDR** - vedi [2]) si ha a disposizione una **base di partenza** nella progettazione che oggettivamente limita la creatività formale-strutturale-funzionale, perché ne impedisce una libera applicazione allo sviluppo del prodotto, **ma** facilita la definizione di forme nuove accettabili e migliorate perché, appunto, evita evoluzioni che si riveleranno molto probabilmente non accettabili se sottoposte ad una analisi di proponibilità sul mercato (analisi da basarsi sulle caratteristiche già richiamate di *user friendliness*) [2]. Questa base è aggiuntiva rispetto ai classici PDR relativi alla funzione, ai *target* di costo e di utente, alla durata e al lotto previsto di produzione, anche se è intrinsecamente legata alla funzione e all'utente e costituisce, quindi, un'enfaticizzazione dell'attenzione su di essi.

E' interessante notare come, a volte, la conservazione dell'interfaccia utente-manufatto richieda un'accurata valutazione della situazione funzionale in studio, per identificare preventivamente eventuali variazioni dell'utenza legate alle nuove funzionalità perseguite (o ritrovate nel corso della progettazione creativa). E' il caso dell'impiego di soluzioni progettuali o parti di soluzioni per altre funzioni o altri scopi: può accadere che il passaggio da una funzionalità ad un'altra stravolga completamente la posizione e la qualità (fisica, psicologica, sensoriale e funzionale) dell'interfaccia utente-manufatto rispetto a quella originaria e che lo stesso utente "cambi configurazione".

Un esempio chiarificatore di quanto sopra riportato è la **poltrona Joe**, ispirata a un guanto da *baseball* in onore di Joe Di Maggio (1971, De Pas, D'Urbino e Lomazzi), fig. 2.



Figura 2. Poltrona Joe. (immagine tratta da [4])

In questo caso evidentemente l'interfaccia utente-manufatto si è spostata da “dentro” a “fuori”, nel passaggio dal modello ispirativo al *concept*, generando una forte alterazione dei *design drivers* (in relazione a dimensioni, consistenza strutturale, materiali, finiture etc.). E' interessante notare, per inciso, che l'idea è stata copiata con variazioni sul tema e utilizzando materiali molto diversi (usando PVC e passando alla soluzione “pneumatica” sono state prodotte poltrone praticamente identiche ad un guantone, con cuciture, colori ed accessori).

In particolare, in questo esempio, il concetto stesso di utente si biforca qualitativamente e quantitativamente, seguendo la duplice funzione del manufatto citato (seduta e oggetto di arredo a vocazione estetica) e negando un'univoca caratterizzazione fisico-sensoriale dell'interfaccia utente-manufatto con l'aggiunta dell'aspetto psicologico in senso lato. Conseguentemente, l'interazione tra manufatto ed utente può (deve!) assumere anche connotati psicologici (che travalicano il puro aspetto pavloviano di controllo, precedentemente accennato, dell'utente durante l'attività dell'unità funzionale), che **devono essere previsti e controllati** (mai subiti!) dal progettista per migliorare efficienza, funzionalità e *customer satisfaction*.

Questa considerazione porta ad una estensibilità (non sempre facile, ma neanche concettualmente complessa) della proposta metodologica al campo del *Design*. A questo aspetto si può ricollegare il fenomeno dello *skeuomorphism* (imitazione priva di necessità funzionale [3]). Questo si manifesta come progettazione e realizzazione di oggetti nuovi con una forma vecchia (consueta, accettata, tradizionale etc.) che permetta di non sorprendere o spaventare il cliente/utente. Lo *skeuomorphism* in senso generale non sempre caratterizza progettualità ottimizzate e razionalizzate, per l'ovvia ragione che se il progredire della tecnologia e sviluppo culturale rende inutili e obsolete (e quindi inefficienti, costose, dannose, impattanti o non sostenibili, perché non ottimizzate!) certe forme-strutture-applicazioni materiche la loro riproposizione deve avere solide motivazioni, che devono travalicare la semplice moda e lo sfoggio di creatività *retrò* (**che è creatività fortemente limitata dai modelli standardizzati**). D'altra parte, una forma-struttura-procedura consueta, che tranquillizza e rassicura l'utente (ovvero genera un'interazione manufatto-utente con una componente psicologica persuasiva ed efficace e contemporaneamente distensiva), può essere usata dal progettista avveduto e responsabile

per ottenere un migliore funzionamento complessivo del sistema manufatto-utente (tenendo d'occhio un obiettivo e ragionevole bilancio costi/benefici) e una conseguente fidelizzazione virtuosa (nel senso di non ottenuta con inganni psicologici opportunistici) del cliente

In conclusione, pare ragionevole (e proattivo) usare come *driver* progettuale nella riprogettazione innovativa di prodotti (o nella *cross-fertilisation* dei *concepts*) **il mantenimento o miglioramento funzionale della configurazione dell'interfaccia utente-manufatto**, sia in senso qualitativo che in senso quantitativo, se l'interfaccia è a livelli **definitivi** di funzionalità. Occorre, pertanto, evitare ogni manomissione dell'interfaccia cioè ogni intervento, deliberato o no, di variazione (anche con scopi estetici o funzionali) che comporti effetti peggiorativi (o non valutabili e/o controllabili con sicurezza) rispetto all'esistente (allo stato dell'arte!).

La sostanziale inerzia su base antropologica (rispetto alla scala temporale di durata dei manufatti) dei comportamenti umani, come identificati da Ergonomia e Psicologia, non permette in questa parte specifica ed essenziale del sistema manufatto-utente l'applicazione di una creatività progettuale assolutamente libera. Inoltre, la pretesa che sia l'utente ad adattarsi *sic et simpliciter* al manufatto (naturalmente sono sempre accettabili - limitate - procedure di addestramento all'uso efficiente) è un inutile (e pericoloso, spesso) sfoggio di mancanza di empatia e di razionalità da parte del PDT. Deve essere applicata la filosofia progettuale *kaizen* che prevede l'innovazione per piccoli passi consolidati nel tempo.

Un esempio è la riprogettazione innovativa di un semplice strumento manuale da taglio (*vulgo* coltello, fig. 3): la configurazione standard dell'impugnatura ha raggiunto livelli di efficienza tali (e contemporaneamente struttura della mano e abilità gestuale dell'uomo non cambiano con velocità rilevabile!) che, nello sviluppo di nuovi prodotti, è impensabile intervenire sostanzialmente su di essa senza peggiorarla: la sagoma riportata in fig. 3 si ritrova in strumenti (e armi) millenari.

In altre situazioni, invece, la possibilità di aggiungere funzionalità nuove e diverse (o la disponibilità di nuove soluzioni per sottofunzioni importanti) al prodotto è in qualche caso collegata ad una diversa e innovativa allocazione e/o configurazione (tecnica, fisica, sensoriale, topologica) dell'interfaccia utente-manufatto, spesso semplicemente aggiuntiva rispetto a quella principale e consolidata. Le alternative funzionali possono essere "esclusive" (per una certa funzionalità solo una certa configurazione) o "non esclusive" (configurazioni opzionali per una data funzionalità), e in qualche caso la configurazione più semplice (e quindi più affidabile) funge da *back-up* di sistema.



Figura 3. Cutter industriale. (immagine tratta da [5])



Figura 4. Trolley bag. (immagine tratta da [6])

Un esempio molto banale in riferimento a questi casi sono le borse dotate di tracolla e maniglia, componenti che svolgono la funzione di vincolo e sostegno da parte dell'utente ma con modalità molto diverse anche in relazione alla funzione globale (generalmente maniglia per “bassa portata” e tracolla per “alta portata”, ma non solo). Le due diverse modalità comportano una diversità sostanziale dell'interfaccia utente-manufatto, con effetti sull'architettura e l'*embodiment* del prodotto stesso (in relazione alla fruibilità di tasche laterali o alla presenza di parti che possano danneggiare gli abiti per sfregamento etc.). Le cosiddette valige *trolley* (un esempio molto semplice è riportato in fig. 4) rappresentano, in questo campo, l'applicazione attuale dei concetti sopra espressi.

3. CONCLUSIONI

La proposta indicazione metaprogettuale di una identificazione conscia ed esplicita di *design drivers* (nella definizione data in questo testo), può semplificare sensibilmente la progettazione, indirizzando il processo rapidamente e con precisione verso soluzioni efficienti ed efficaci, se recepita fino dalle prime fasi dell'*iter* progettuale.

Si ribadisce, a questo proposito, il concetto già espresso relativo alla limitazione “virtuosa” della libertà creativa progettuale in vista di una funzionalità più affidabile e sicura e di una sopravvivenza prolungata del prodotto sul mercato.

I due *design drivers* identificati per la progettazione di prodotti ad alta interazione operativa utente-manufatto (integrazione funzionale utente-prodotto e mantenimento o miglioramento funzionale dell'interfaccia utente-manufatto), al di là dei semplici esempi, costituiscono una solida base per l'identificazione di nuovi *concepts*, e per i successivi *embodiment* di prodotto, funzionali e sicuri.

BIBLIOGRAFIA

- [1] Pahl G., Beitz W, 1984. Engineering Design. Springer-Verlag, Berlin.
- [2] Medri G, 2008. Cenni di Progettazione di Prodotto. Liguori ed., Napoli.
- [3] Molotch H., 2005. Fenomenologia del tostapane. Cortina ed., Milano.
- [4] <http://www.designindex.it>.
- [5] <http://www.vola-instruments.com>.
- [6] <http://www.biz4cn.cn/productdetail/7709/27420/Bags>.

PROJECT-BASED CO-TUTORED MECHATRONIC LEARNING

Gabriele Vassura
*DIEM -Department of Mechanical Engineering,
University of Bologna, Italy
E-mail: gabriele.vassura@unibo.it*

Alessandro Macchelli
*DEIS – Department of Electronics and Information Systems
University of Bologna, Italy
E-mail: amacchelli@deis.unibo.it*

Abstract. *A project-based learning experience has been recently developed at the University of Bologna within the second-level degree in Automation Engineering. The most relevant aspect is a co-tutoring activity jointly performed by teachers from the mechanical area and from the automatic control area. The project consists in the design of an automated assembly system, developed for a local company leader in the production of technical cases. After a description of the educational goals, the paper discusses the phases of the activity and the main methodological aspects, briefly presents the adopted tools for the design and simulation of the developed system and finally discusses the achieved results..*

Keywords: *Education, automation, mechatronics, project-based learning, multidisciplinary tutoring.*

1. INTRODUCTION

In this introduction will be explained the general considerations and the local needs that led to the institution of a curriculum in Automation Engineering at the University of Bologna, Italy, and, inside this frame, to the organization of a project-based co-tutored learning activity aiming to the development of professional engineering skills in automation and mechatronic design.

Automation and Mechatronics

It has been widely remarked in the literature, e.g. [1], that Automation and Mechatronics domains are only partially overlapping and the two words are not strictly synonyms; indeed, many automation solutions cannot be considered mechatronic, as well as many mechatronic systems are not automated systems. No doubt, however, that modern flexible automation is more and more applying mechatronic solutions and that an automation engineering curriculum cannot be developed without the explicit or implicit

proposal of a mechatronic approach to design. For this reason the definition of mechatronic learning was used in this case, even if the relevant aspect of the activity is not strictly focused on mechatronic product development, but rather on the synthesis of an automated complex system where mechatronic concepts have been applied.

Industrial needs and engineering education

The industrial area around Bologna is known worldwide for the high density of companies involved in the production of automated systems, especially of packaging machines. Most of these companies were founded by ingenious technicians very skilled in mechanical design and became world leaders thanks to their capability to provide a quick and top-level response to a market characterized by extreme customization.

From a few mother-companies a myriad of small enterprises spread out, thus providing a dissemination of technical knowledge and practical know-how and demonstrating the capacity to operate with highly specialized technology as well as organizational flexibility.

Inside these companies design roles were covered at the beginning mainly by people coming from a practical mechanical education, at the best with high-school degree or simply with machine-shop experience, only later by mechanical engineers. Electronic and information technology skills became progressively necessary with the development of modern technologies (in particular with the massive advent of electromechanical drives and programmable controllers) but, in most cases, they were integrated into design teams with ancillary functions, seldom participating to the early conceptual design of the automatic machine. Nowadays too in most of the automatic machinery companies the initial design is performed mainly by people with dominant mechanical background.

It is widely demonstrated that in most cases such an approach cannot fully exploit the potential that rising mechatronic technologies can provide. In particular, there are valid evidences that people with a mono-disciplinary education, even if open-minded and well informed about the complementary available technologies, tend to apply to problem solving (and therefore to the development of innovation) a point of view that is heavily conditioned by their dominating cultural background.

Automation Engineering First-level degree (3 years)		
First year	credits	
• Mathematical Analysis *	15	
• Geometry and Algebra	6	
• Physics I	9	
• Found. of Computer Science *	12	
• Found. of Mechanics	6	
• Found. of Industrial Economy	6	
• Foreign Language	3	
Second year	credits	
• Physics II	6	
• Electrotechnics	6	
• Found. of Electronics	12	
• Automatic Control *	18	
• Applied Mechanics*	18	
Third year	credits	
• Electromechanical drives *	12	
• Computer technology	6	
• Control system technology	6	
• Automatic Machines	6	
• Industrial Robotics	6	
• eligible courses **	12	
• eligible laboratory activity***	12	
• final work	3	

* Articulated in 2 separate didactical modules
 ** Available 11 courses shared with other curricula
 *** Available Workshops on Robotics, Process Automation, Computational Modelling

Figure 1. Educational activities in the three-year first level curriculum

In order to overcome the limits of traditional engineering curricula, new educational profiles are under definition and test world-wide, according to the shared opinion that a balanced multi-disciplinary education can help to approach in a more creative way design problems that involve the integration of so many technologies. Mechatronics is rising to the role of a self-standing discipline whereas education in mechatronics is considered a very specific goal worthy to be achieved. In some cases the new curricula are explicitly called Mechatronic Engineering, in other cases the denomination of Automation Engineering is preferred, but in general small differences can be found between the contents that characterize the two curricula.

Automation Engineering curriculum at the University of Bologna

In Italian engineering schools, in particular at the University of Bologna, an educational curriculum explicitly oriented to integration and synergy of different cultural domains (mechanics, electronics, information science) has been proposed only in recent years. A slowing factor, apart the difficulty to harmonize contributions from academic partners traditionally not trained to cooperate, were the many doubts about the consistence of such a professional figure. Indeed, the mechatronic engineer risks to be one who owns knowledge in multiple fields, but has no actual operating capacity in any of them. Further doubts were arising about the capacity of the industrial world to be ready to accept and exploit such a professional figure. After a period of careful evaluation, a new curriculum in “Automation Engineering” was defined and its implementation started.

In Figure1 a scheme of the basic three-year educational activities is presented in its latest revision. It can be seen that, since the first degree level, courses characterized by multidisciplinary integration (e.g. Industrial Robotics and Automatic Machines) are available and represent a consistent part of the didactical offer. Activities oriented to design however are not significant (laboratory activities mainly consist in demonstration of basic components and instruments) and emphasis is put mainly on the acquisition of basic knowledge in the different disciplinary areas.

Automation Engineering Second-level degree (2 years)			
First year	credits	Second year	credits
• Adv. Mathematics	6	• WS Automatic Machinery	6
• Adv. Automatic Control *	6	• WS System Automation	6
• Operational Research	6	• Adv. Ind. Electronics	6
• Adv. System Theory	6	• Innovation Management	6
• Distributed Control Systems	6	• Electromagn. Compatibility	6
• Robot Dynamics	6	• eligible courses*	12
• Adv. Computer Technology	6	• eligible laboratory activity	6
• Production Process Technology	6	• Industrial stage **	6
• Adv. Logistics	6	• Final thesis work	6

* Available 8 courses shared with other curricula

!!! Curriculum under restructuring

Figure 2. Educational activities in the two-year second-level curriculum

The activities planned for the second level degree (*laurea magistralis*) in the first year offer a balanced mix of advanced contributions in mechanics, information science and electronics in order to integrate the basic knowledge provided in the three- year curriculum in the second year are explicitly oriented to encourage the growth of design capability, proposing consistent learning-by-doing activities.

To this purpose the proposed workshop activities are focused on project development, mainly working on case-studies directly suggested by industrial partners.

This paper reports about the start-up phase of such a learning-by-doing activity, resulting from the coordination of two educational modules, one mainly focused on mechanical engineering aspects (Automatic Machinery Workshop) and the other on control engineering aspects (Automation Systems Workshop), in the attempt to develop a pragmatic and hopefully efficient experience of education to mechatronic design.

2. THE PROJECT-BASED LEARNING ACTIVITY

Educational goals of the activity

The main educational goals of the joined workshop activities can be summarized as follows:

- to stimulate the creativity of the students involving them in the solution of a design problem of real interest for industry, never faced before, needing a synergetic multidisciplinary approach; emphasis is put on the conceptual design and on the organized comparison of the generated design variants, in order to demonstrate the potential of a mechatronic approach with respect to different approaches;
- to provide them methodological indications on how to move along the process of development of a complex automated system;
- to achieve specific skills in the use of design and simulation tools, jointly used in the development of the different aspects of the project;
- to train their capability to analyze and describe a problem, to generate and evaluate solutions under a concurrent set of design constraints, to present and defend the results of their work, not neglecting a critical review of the achieved results.



Figure 3. The dimensional family of Explorer cases (courtesy of GT Line)

The choice of the project theme

The choice of a proper project theme is in general a very critical matter, as it must cope with numerous requirements:

- to be well calibrated as to complexity, avoiding to be trivial or too difficult with respect to the student skills; furthermore, its development time must be compatible with the allowed activity period (about twelve weeks); the development charge must be compatible with the correspondent credit assignation;
- to require synergetic application of differentiated disciplinary knowledge, offering the opportunity to apply a methodological approach to design and to possibly involve all the aspects of mechatronic design;
- to offer good novelty level, so that conceptual design can be developed freely, without suggestions coming from existing solutions;
- to be a theme of industrial interest, with a defined industrial partner, so that both the acquisition of design specifications and the verification of results can be done in a realistic context;
- to be stimulating for the students, in order to obtain the best from their creativity;
- to be compatible with the supervisors' skills.

A suitable project theme was identified in the automatic assembly of retaining hooks on the top cover of technical cases. These cases, made with reinforced moulded polypropylene, are produced in different sizes and models by GT Line, a market leader company operating in Bologna (Figure 3).

At present the assembly process is completely manual and represents a consistent part of the final cost. A cooperation between GT Line and Mechanical Engineering Department DIEM for the development of an integrated assembly system of cases was active since about one year, so that many aspects of the problem had already been evaluated and the overall feasibility examined. The design activity of the students was limited to the development of a sub-system, required to pre-assemble the retaining hooks (three parts) and insert a variable number of such hooks onto the cover of each case (every hook must be positioned with respect to the cover edge, then a stainless steel pin must be laterally inserted).

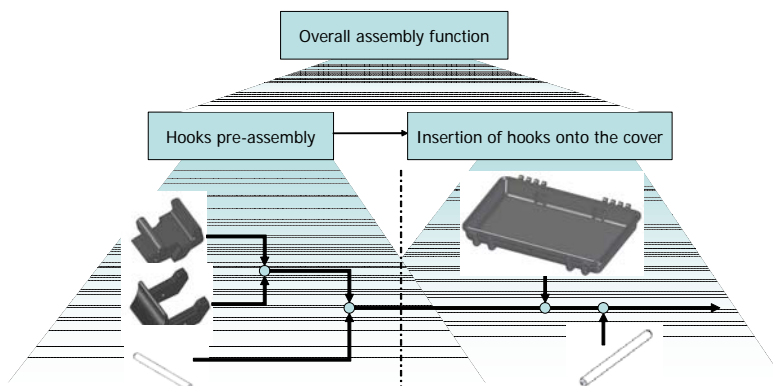


Figure 4. The assembly tree of retaining hooks and their insertion on the cover

The hook is the same for the whole family of cases, but must be positioned differently according to the size of the cover, therefore intrinsic flexibility of the insertion system is required, because the batch-production mix can be extremely variable and covers with different size and position of hooks can alternate on the assembly station.

The overall assembly function was analysed and subdivided into two separate assembly problems (see Figure 4) that were each examined by a single group of students, with continuous exchange of information in order to check the compatibility of the developed sub-systems.

Phases of the project and adopted CAD tools

The activity has been developed both as class-work (in total 6+6 hours per week for about 12 weeks) and as home-work (an estimated average amount of 20 hours per week). The students were 12 in total and were divided in two cooperating subgroups. Each subgroup developed detailed design of a part of the automated system, but the initial activity of the conceptual design of the whole system was developed jointly.

In Figure 5 the main steps of the overall activity are presented.

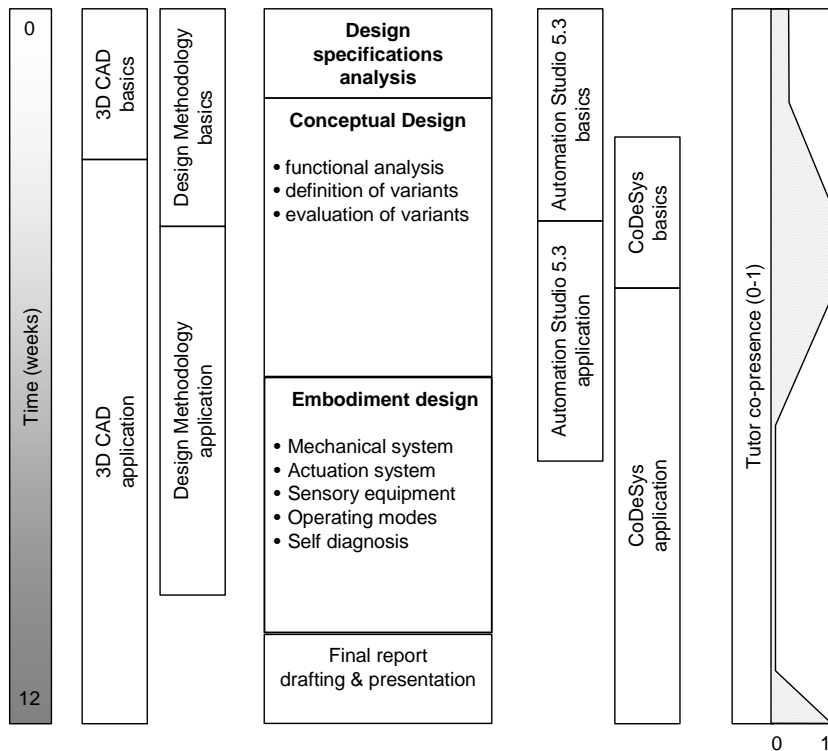


Figure 5. Main steps of the activity and adopted CAD tools

The following remarks can be added:

- in parallel with the analysis and discussion of the possible solutions to the specific automation problem, a consistent part of the class-work was dedicated to the acquisition of skills with important design tools, a 3D Cad (Solid Edge) for physical description of the designed system, a general tool for virtual prototyping of the automated system (Automation Studio 5.3) and a purpose PLC simulation tool for a more detailed analysis of the control sequence of the system (Codesys);
- the development of the system was really based on a mechatronic approach, that means that all the aspects concerning actuation, sensory equipment and control were considered jointly with the design of the mechanical part of the system and a sort of simultaneous engineering was applied;
- the students were required of a final public presentation of their work and attention was paid on the efficacy of communication; the industrial partner was asked to interact with the students during this presentation, with an explicit evaluation of the proposed technical solutions.

Tutoring aspects

Tutoring has been intended as a resource available to students in order to be assisted in the design process. One tutor was coming from the area of mechanical design, the other from the area of electronics and automatic control, so that experienced skill for most of the aspects of mechatronic design could be provided.

The two tutors most of the time were operating inside own time slots, but were present together in all those phases of the group-work (conceptual design) that required joint evaluation of the proposed technical solutions, system architecture and related operating sequence, as well as in the final phases when the students had to prepare the presentation of their results. A qualitative representation of the level of co-presence of the tutors at the group work is sketched in Figure 5, right. It can be seen that, after the definition of the overall conceptual design of the system, embodiment design activities were mainly tutored separately.

The role of tutors was different in the various phases of activity. At the beginning this role was active and consisted mainly in the proposal of methodological guidelines for the development of the project and in an organized presentation of the tools to use in analysis and simulation. As soon as the students started the proposal phase, the function of the tutors gradually changed and their main effort was to stimulate their creativity, suggesting correct ways to examine the problems and generate solutions, helping to organise but not conditioning their decisional process.

The “mechanical” tutor first provided the basics of methodological design, in order to organize the decisional process of design according to a define-evaluate-compare-choose sequence, avoiding the main temptation of beginners to follow a decisional process based on local intuitions. To this purpose, the basic steps for the development conceptual design according to VDI 2225 recommendations [2] were illustrated and substantially followed in the development of the project, with particular emphasis on functional analysis. To better understand the technical problems and the economical constraints and define the design specifications, the students spent some hours inside the company, analyzing in detail the manual assembly process and the actual organization of the workplaces.

In parallel, basics of 3D CAD representation, missing in the students background, were taught: the capability to represent the result of the inventive work was considered an

essential pre-condition before starting the practical development of the project. A learning-by-doing procedure was imposed by the lack of time, but after a few lectures about the basics of 3D representation, the students were able to start working, getting progressively skilled and able to represent the evolution of their virtual prototype.

The parallel work of the “electronic” tutor further developed the skills related to the general organization of mechatronic control systems: the students had good previous knowledge of the basic automation components (sensors, pneumatic actuators, electric drives, etc.) and of the main standards for design and representation of work sequences of automated systems (e.g. IEC61131-3 and IEC61499) but were lacking of experience in the use of advanced modular tools for development of software for industrial automation. An object-oriented modelling method, based on the definition of a “generalized actuator” [3,4] and implemented in CoDeSys environment was chosen; preliminary design of the automation system was verified with the help of Automation Studio 5.3, but final development and verification of the PLC software was performed inside CoDeSys [5].

Results

The final result of the design developed by the students is shown in Figure 6., 7 and 8. In order to obtain the required flexibility, the task of manipulating the covers has been given to a robot, for which a special purpose gripping device based on suction-cups has been designed. The robot is interfaced with a stationary equipment, with two terminal stations where two or more hooks can be positioned. This duplication is necessary due to the fact that the pins of the hooks placed near the corners of the cover need to be inserted from opposite directions.

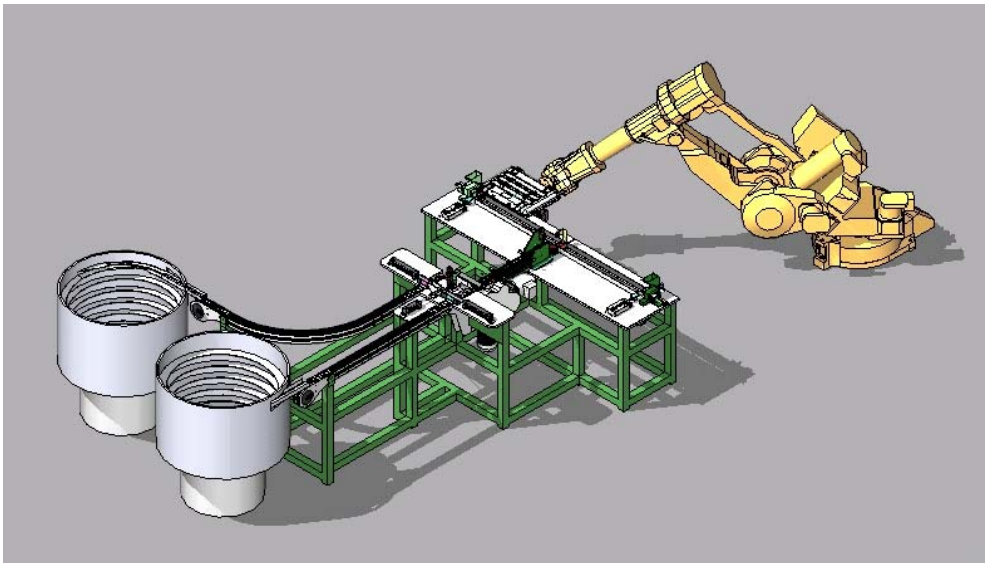


Figure 6. Overall view of the developed assembly system

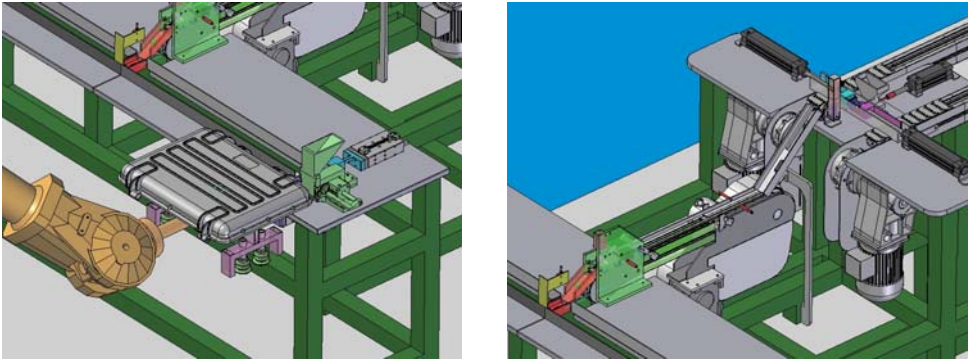


Figure 7. Some more details of the designed automated system

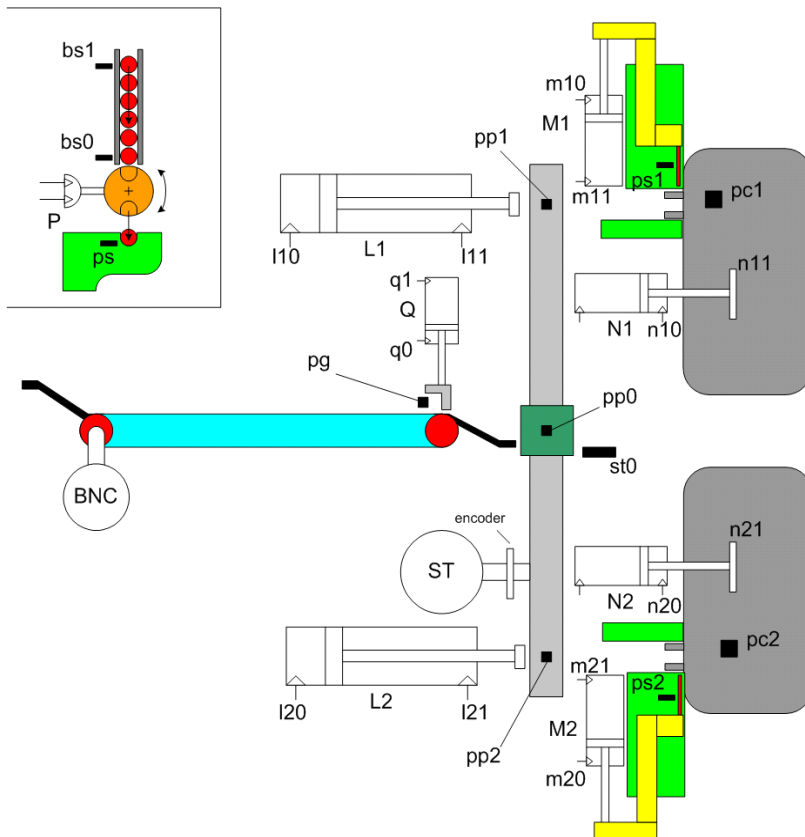


Figure 8. A model of the inserting station for the preliminary analysis of the operating sequence (using Automation Studio 5.3)

Upstream with respect to the two positioning stations is placed a system that, starting from parts fed in bins, orients and mates the parts constituting each hook. Pre-assembly and inserting stations are distant, so that intermediate manipulation and transfer operations are needed; the adopted belt feeders perform also buffering functions, in order to add operating flexibility to the system.

Actuation, sensory equipment and control apparatus have been chosen depending on the functional requirements; particular attention was paid to include self-diagnosis capability, obtained by additional sensors and by purposely defined control procedures. To this respect the simultaneous design of the physical part of the system and the development of control and diagnosis procedures was very influent on final design. The developed system does not adopt very sophisticated technology, but is a complete mix of commonly used solutions (pneumatics, vacuum technique, position-controlled axes, etc.).

Comments

Globally speaking, the authors consider this experience successful and worthy to be repeated: besides the tutors' opinion, very positive evaluation was expressed by the protagonists, the students, and by the industrial partner.

The success of this didactical experiment depended on several factors:

- a consistent, fairly balanced background achieved by the students in the basic disciplines that contribute to the mechatronic domain; hardly such an activity could be proposed to students at the early stages of their technical education;
- a proper choice of the project to be developed, that presented a good level of difficulty, equally distributed between mecha- and -tronic areas, but not, at the same time, too complex;
- the parallel simultaneous development of all the design aspects related to mechatronic design, overcoming the limits of traditional approach (where the mechanical engineer conceives and designs a system, then passes the project to a control engineer that must make it move); efficient planning of this joined activity was a source of problems, but we think it was preferable with respect to a serial concatenation;
- the enthusiastic approach by the students, strongly stimulated by the responsibility to be autonomous managers of their work and by the trust that the company showed respect their creativity and technical capacity;
- a quasi full-time activity, with great attention to the communication inside the workgroup and to the joint participation of students and tutors to the generation of ideas and solutions;

Among the limits, the incompleteness of the project, that was limited to strictly technical aspects and could not develop, due to lack of time, a more detailed evaluation of economical aspects. One of the most challenging perspectives could be to integrate the contribution of a third tutor, purposely operating on this crucial aspect.

No doubt that the attempt was successful also due to the limited number of students involved in this start-up year: this fact allowed a very capillary interaction between the tutors and each participant, favouring a homogeneous level of personal involvement and equally subdividing the charge among the group components.

3. CONCLUSION

The authors focused their efforts on the attempt to develop mechatronic design capabilities according to a methodological approach. Well tested, comprehensive methodologies for mechatronic design of complex technical systems are still under development: more than proposing design rules mechatronics literature so far has mainly proposed case-studies. The adopted tools were, on one side, the general guidelines for systematic design of technical systems, with a key role of functional analysis and variant generation methods (like the morphological matrix) that allowed to efficiently identify and compare possible design variants, more or less inspired to mechatronic concepts. This greatly helped the students to understand the potential of mechatronic approach with respect to more traditional design. On the other side, new tools for virtual prototyping of computer controlled systems, that allowed an efficient development and verification of the functional capabilities of the generated concurrent proposals. Both methodological tools proved to be efficient and useful, but probably a more explicit integration between them is still to be developed: co-tutoring was very important in order to overcome in practice this limitation. From their experience the authors cannot yet define any general rule or recommendation, but are conscious that the attempt provided good results and is worthy to be repeated: next challenge will be to transform this experiment into a consolidated procedure, making the result not dependent on the start-up favourable conditions (e.g. the reduced number of students).

REFERENCES

- [1] Isermann, R., 2005. *Mechatronic Systems Fundamentals*, Springer, London, ISBN 1-85233-930-6
- [2] Pahl G., Beitz W., 1999. *Engineering Design; A Systematic Approach*, Springer, London, ISBN 3-540-19917-9
- [3] Bonfé M., Fantuzzi C., Secchi C., 2006. "Behavioural inheritance in object-oriented models for mechatronic systems". *International Journal of Manufacturing Research*, **1**(4), pp. 421–441.
- [4] Faldella E., Paoli A., Sartini M., Tilli A., 2008. "Hierarchical supervision systems in industrial automation: a design procedure based on the generalized actuator concept". *17th IFAC World Congress*, Seoul.
- [5] Vassura G., Macchelli A., 2008. "Multi-disciplinary Tutoring for Project-Based Mechatronics Learning" *17th IFAC World Congress*, Seoul, invited session on Mechatronics Education.

STUDIO NUMERICO E SPERIMENTALE DEL COMPORTAMENTO MECCANICO DI LASTRE IN ACCIAIO ED IN ACCIAIO PORCELLANATO SOGGETTE AD IMPATTO A BASSA VELOCITA'

Andrea Zucchelli

Dipartimento delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia (DIEM),

Facoltà di Ingegneria, Università di Bologna, Italia

E-mail: a.zucchelli@unibo.it

Giangiaco Minak

Dipartimento delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia (DIEM),

Facoltà di Ingegneria, Università di Bologna, Italia

E-mail: giangiaco.minak@unibo.it

Daniele Ghelli

Dipartimento delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia (DIEM),

Facoltà di Ingegneria, Università di Bologna, Italia

E-mail: daniele.ghelli@unibo.it

Abstract. *Scopo del lavoro è la caratterizzazione del comportamento ad impatto di lamiere rivestite di smalto porcellanato dal punto di vista sperimentale e lo sviluppo di un modello agli elementi finiti che consenta di valutare gli effetti di diversi parametri di processo, tra cui le tensioni residue indotte dalla differenza dei coefficienti di dilatazione termica di substrato e rivestimento. Il medesimo modello è stato utilizzato per simulare l'insorgenza di un particolare fenomeno locale di instabilità dell'equilibrio.*

Keywords: *impatto a bassa velocità, smalto porcellanato, instabilità, simulazioni numeriche*

1. INTRODUZIONE

I rivestimenti ceramico-vetroso, noti anche come rivestimenti porcellanati, possono essere applicati su matrici ceramiche, sulla terracotta, sui metalli e sui vetri. L'attività di ricerca che è stata sviluppata in merito a tale tipologia di materiali ha riguardato i rivestimenti ceramico-vetroso per substrati metallici, noti anche come "smalti porcellanati". L'origine di tali rivestimenti si perde nei millenni (le prime testimonianze di tali rivestimenti sono riscontrabili in applicazioni su gioielli e monili di epoca egiziana.), e sono quasi sempre stati apprezzati per l'aspetto esteticamente accattivante e per la loro capacità di rendere i substrati totalmente impermeabili all'acqua. Solo recentemente (XX

Secolo) il loro utilizzo è stato esteso ad applicazioni tecniche per incrementare la resistenza chimica di alcuni manufatti quali i riscaldatori d'acqua o gli elementi degli scambiatori di calore delle centrali termoelettriche, nonché per conferire maggiore pulibilità e stabilità della finitura superficiale a manufatti finalizzati a molteplici impieghi industriali [3-5] (pannelli per l'edilizia pubblica e componenti meccanici per uso biomedicale). Gli smalti porcellanati sono di natura inorganica e la loro struttura interna, nella quale vi sono anche inclusioni bollose, è caratterizzata dalla presenza di una prevalente fase vetrosa (superiore all'85%) e da una forte integrazione con il substrato metallico. In particolare l'integrazione del rivestimento con il substrato metallico avviene grazie alla presenza di una zona di interfaccia caratterizzata da una struttura che ospita delle dendriti metalliche e che varia in modo graduale la sua composizione (functionally graded bonding interlayer). Tale interfaccia è una diretta conseguenza del processo di smaltatura che avviene ad elevata temperatura (variabile a seconda del tipo di metallo; in particolare per l'acciaio è di circa 870°, per la ghisa è di circa 760° mentre per l'alluminio è prossima a 610°C) ed è costituito da un materiale ibrido amorfo nel quale il metallo si combina con alcuni ossidi che costituiscono il rivestimento e gradualmente in esso scompare. In virtù di questa caratteristica lo smalto porcellanato viene a formare con il substrato metallico un "corpo unico" ed in tal senso si può parlare di "metallo porcellanato". Altro aspetto che caratterizza l'acciaio porcellanato è lo stato di tensioni interne che si sviluppa durante il processo produttivo. Infatti lo smalto porcellanato ha un coefficiente di dilatazione termica inferiore a quello del substrato metallico e questo, durante il raffreddamento, causa l'instaurarsi di tensioni di compressione nello strato ceramico-vetroso e di trazione nel metallo [2]. Questo stato di tensioni residue che esistono nel metallo porcellanato ne caratterizza il comportamento meccanico sia quando il componente con esso realizzato è soggetto a carichi quasi statici, sia quando è soggetto a carichi di impatto a bassa velocità come verrà di seguito discusso. Nel lavoro che viene qui presentato viene posto a confronto il comportamento meccanico di lamiere in acciaio e di lamiere in acciaio porcellanato soggette ad un impatto a bassa velocità.

Nell'ambito dello studio sono state considerate lastre rettangolari sia in acciaio che in acciaio porcellanato e per entrambe le tipologie di provini è stata evidenziata la presenza di una instabilità localizzata sui bordi dei provini stessi, Figura 1, quando l'energia di impatto supera una data soglia limite. Dallo studio sperimentale si è potuto notare che nel caso delle lamiere in acciaio porcellanato l'energia d'impatto necessaria per innescare l'instabilità risulta essere sei volte maggiore rispetto al caso delle lamiere in acciaio. Lo studio numerico del fenomeno ha messo in evidenza che l'instabilità è in generale associata ad un plasticizzazione del materiale nell'area ove si verificano i maggiori spostamenti dei bordi liberi. Inoltre è stato possibile verificare che il carico massimo che viene registrato durante le prove sperimentali di impatto nelle quali si è verificata l'instabilità è prossimo al valore del carico che nel corso delle simulazioni numeriche causa l'innescare della instabilità. Infine mediante i modelli numerici è stato possibile studiare l'influenza delle tensioni residue sul carico che causa l'innescare della instabilità.

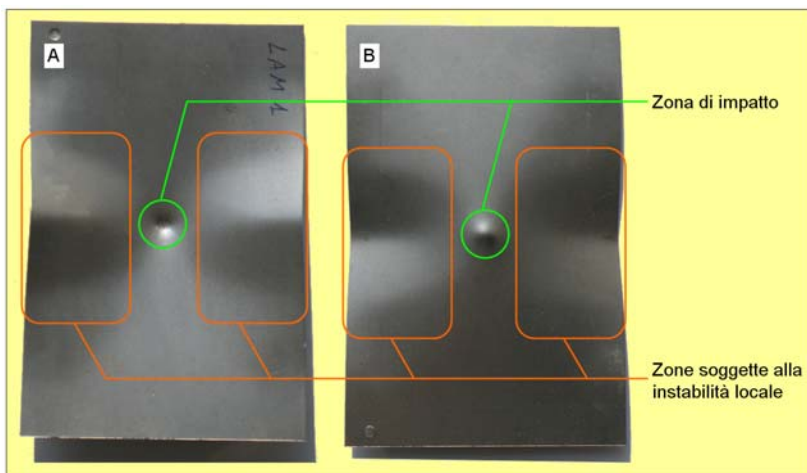


Figura 1. esempio di lamiera in acciaio nella quale si è verificata l'instabilità a seguito di urto a bassa velocità: (A) fronte della lamiera, (B) parte posteriore delle lamiera

2. MATERIALI E METODI

Materiali e metodi sperimentali

I materiali utilizzati per le prove sono lamiere di forma rettangolare (150mm x 100mm) realizzate sia in acciaio che in acciaio porcellanato. La lamiera di base ha spessore 0.8 mm mentre per i rivestimenti è stato ottenuto uno spessore medio di 200 μ m per parte. L'acciaio utilizzato appartiene alla classe dei decarburati (DC04ED, ultra low carbon steel) poiché è quello che meglio si presta per ottenere la formazione del rivestimento porcellanato e che viene comunemente utilizzato nella pratica industriale per realizzare gli elementi scambianti degli scambiatori di calore rotativi. I rivestimenti sono stati realizzati mediante smalti porcellanati tipicamente utilizzati per la protezione di componenti di scambiatori di calore per le centrali termoelettriche. In particolare i rivestimenti hanno subito consolidamento termico in forno radiante industriale ad una temperatura massima di 870°C mantenuta per un tempo di 6'30''.

Le prove di impatto a bassa velocità sono state eseguite con una macchina drop-weight considerando differenti livelli di energia nominale (da 3J a 36J). In particolare è stato utilizzato un impattatore avente massa di 1.22 kg, strumentato con una cella di carico piezoelettrica per la misura della forza di contatto. La macchina di prova è dotata anche di un sensore laser per la misura della velocità effettiva dell'impattatore poco prima dell'urto sui provini.

Le piastre sono state fissate su una base di appoggio realizzata in conformità alla normativa ASTM [9] per gli impatti a bassa velocità su compositi polimerici, Figura 2. È stato seguito il metodo suggerito dalla stessa norma anche per il calcolo dello spostamento dell'impattatore mediante integrazione del segnale della cella di carico e per la stima dell'energia assorbita durante l'impatto.

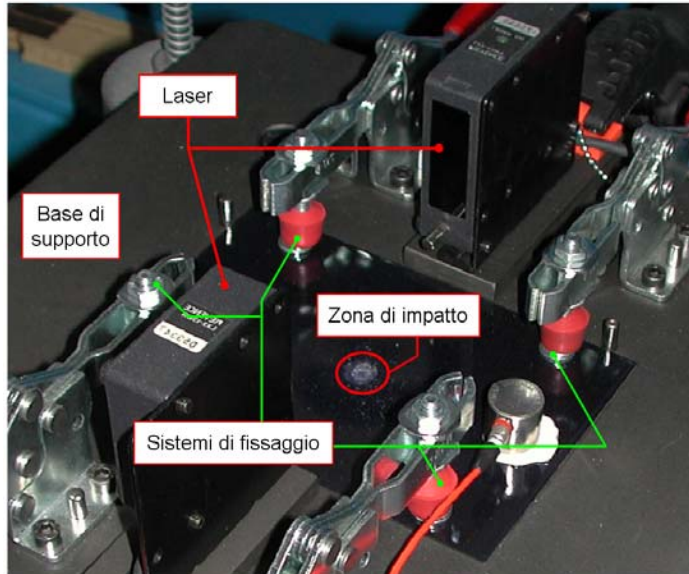


Figura 2. Immagine relativa al sistema di fissaggio dei provini soggetti ad impatto

Tabella 1: piano delle prove di impatto

Campione	Parametri						
Lamiera in acciaio	Altezza [m]	0.25	0.50	1.00	1.50	2.00	\
	Energia [J]	3	6	12	18	24	\
Lamiera porcellanata	Altezza [m]	\	0.50	1.00	1.50	2.00	2.00
	Energia [J]	\	6	12	18	24	36

Il piano sperimentale è stato articolato effettuando prove di caduta dell'impattatore su lamiera porcellanate e non. Le prime prove sono state condotte sulle lamiere non smaltate e sulla base dei risultati ottenuti è stato perfezionato il piano prove per le lamiere smaltate. In tal modo è risultato un piano prove, Tabella 1, che è leggermente differente tra le lamiere non smaltate e quelle porcellanate. La prova indicata con energia pari a 36 J è stata effettuata da una altezza pari a 2.00m aumentando opportunamente la massa dell'impattatore.

Per quanto attiene alle proprietà caratterizzanti utilizzate nel presente studio distinguiamo tra le proprietà fisiche, il coefficiente di dilatazione termico, e quelle più strettamente meccaniche, Tabella 2.

Tabella 2: proprietà dei materiali utilizzati

Materiale	α [$10^{-5} \text{ }^\circ\text{C}^{-1}$]	E [MPa]	ν	σ_y [MPa]	E_t [MPa]
Acciaio DC04ED	1.30	180000	0.3	100	1200
Smalto porcellanato	1.18	70000	0.27	\	\

In particolare le caratteristiche meccaniche dell'acciaio sono state valutate sperimentalmente mediante prove di trazione, eseguite in controllo di spostamento, in conformità alla norma UNI EN 10002, montando due estensimetri nella zona centrale del provino ad osso di cane in direzione rispettivamente longitudinale e trasversale.

Per la lamiera in acciaio, in Tabella 2, sono riportati i dati relativi al modulo di Young ed al coefficiente di Poisson, alla tensione di snervamento ed al modulo tangenziale (E_t) riferendo in tal modo il comportamento del materiale ad un modello inelastico bilineare. I valori relativi al coefficiente di dilatazione termica lineare dell'acciaio sono stati tratti dalla letteratura del settore [1]. Il valore in Tabella 2 relativo al coefficiente di dilatazione termica dello smalto porcellanato è stato ottenuto mediante prove sperimentali effettuate presso il laboratorio della COLOROBIA S.p.A. La misura è stata condotta ponendo una barretta di smalto porcellanato puro (ovvero non su supporto metallico) in un dilatometro riscaldato ed il valore di α è stato stimato mediante interpolazione lineare della curva di dilatazione termica su un intervallo termico di 400°C , Figura 3. Nella medesima figura sono visibili quattro immagini ricavate mediante il microscopio riscaldante nelle quali viene evidenziato il momento in cui lo smalto è ancora nello stato solido (A), mentre nei particolari (B), (C) e (D) sono visibili le zone di fusione progressiva dello smalto in prossimità della zona di appoggio. Si osserva che il valore di dilatazione termica per lo smalto porcellanato (α) è leggermente inferiore a quello della lamiera di base. Questa piccola ma significativa differenza fa sì che la lamiera porcellanata sia soggetta ad uno stato di tensione interna dovuto al ciclo di maturazione termica.

In particolare si osserva che durante il raffreddamento a partire dalla solidificazione, che corrisponde alla temperatura di cambio pendenza (break rate temperature, BRT) prossima a 400°C , lo smalto, a causa del progressivo raffreddamento, ha contrazioni dimensionali inferiori a quelle che segue la lamiera. Questo fatto induce nel composito finale uno stato di tensione residua di compressione nel rivestimento e di trazione nel substrato metallico.

Lo studio dettagliato relativo alle tensioni residue che nascono nel composito porcellanato non è oggetto centrale del presente lavoro, tuttavia verranno prese in considerazione nel corso delle analisi numeriche per studiare il loro effetto sulla risposta complessiva della lamiera soggetta all'impatto.

Per quanto attiene alle proprietà meccaniche dello smalto si è fatto riferimento alle prove di trazione eseguite in occasione dello studio riportato nel lavoro [5].

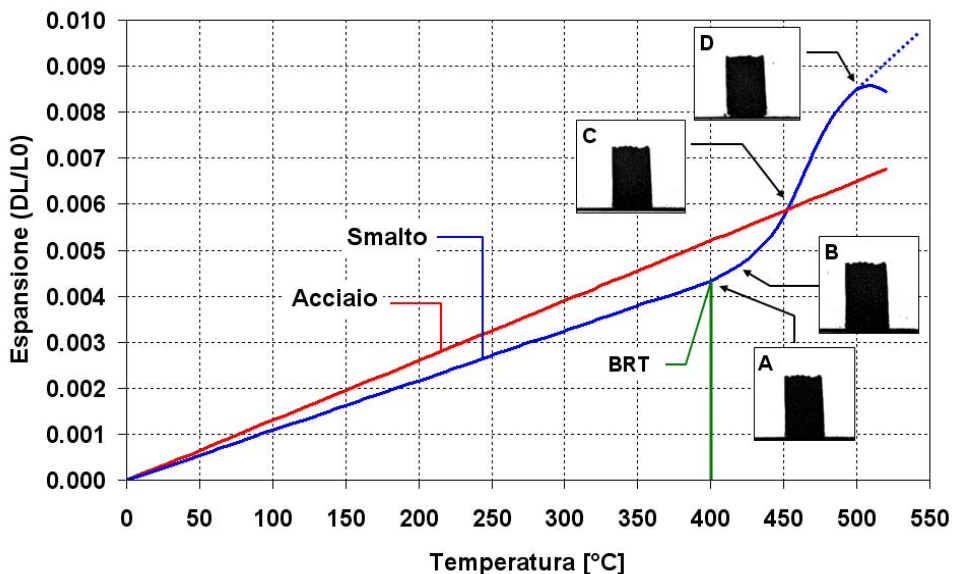


Figura 3. Andamento della espansione lineare della lamiera e dello smalto in funzione della temperatura, con particolari presi al microscopio riscaldante relativi al comportamento dello smalto.

In generale si osserva che dal punto di vista meccanico lo smalto porcellanato è un materiale fragile per il quale non è possibile identificare una soglia di snervamento mentre si può definire un valore limite di tensione al quale avviene la sua rottura. Sulla base delle precedenti esperienze, per lo smalto considerato nel presente studio, è possibile assumere un limite massimo di tensione a rottura pari a 35 MPa. Tale valore limite è stato utilizzato nelle simulazioni numeriche per poter rappresentare in modo realistico la rottura del rivestimento durante il processo di impatto.

Modelli numerici

Le analisi numeriche sono state condotte riproducendo le condizioni di impatto con un carico quasi statico applicato al centro del provino, in direzione trasversale al piano del provino stesso, mentre sono attualmente in corso di perfezionamento le simulazioni numeriche in condizioni dinamiche.

Le analisi in condizioni di carico quasi statico, sono state effettuate mediante un codice di calcolo commerciale, ANSYS, assumendo il comportamento del substrato metallico come non lineare secondo il modello bilineare. Data la simmetria del problema il modello geometrico di riferimento è il quarto di lastra ed il relativo supporto di base.

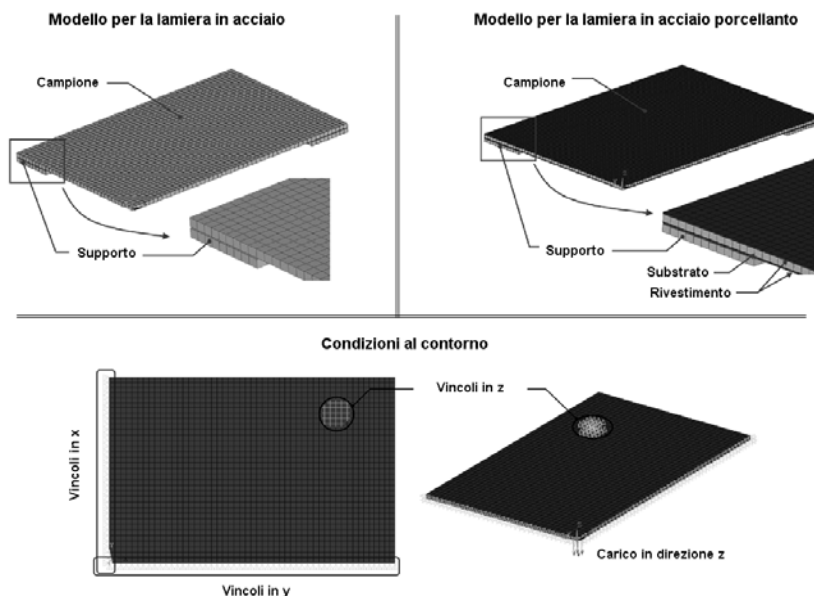


Figura 4. Modello numerico relativo alla piastra in acciaio ed in acciaio porcellanato e condizioni di vincolo.

Il modello numerico è stato ottenuto utilizzando elementi solidi ad 8 nodi per la lastra, smaltata e non, e per la base di appoggio.

La dimensione media degli elementi per effettuare la discretizzazione è stata scelta pari a 1.5 mm. Secondo tale impostazione è stata ottenuta una discretizzazione che consta di uno strato di elementi per la lastra in acciaio ed un singolo strato di elementi per ciascuno strato di smalto.

La zona di contatto tra la lastra e la base è stata simulata utilizzando il metodo del contatto implementato in ANSYS [11]. In particolare sono stati utilizzati elementi contact associati alle aree di contatto della piastra ed elementi target per le aree associate al supporto. La tipologia di contatto è standard mentre per il coefficiente di attrito è stato assunto il valore di 0.7. Infine si fa notare che nella realizzazione del modello tra la lastra e la base è stato volutamente lasciato un interstizio pari a 0.1mm. Per poter stabilizzare la soluzione del modello si è quindi reso necessario imporre come condizione operativa degli elementi di contatto la chiusura del gap e la contestuale riduzione della penetrazione.

Per quanto attiene alle condizioni al contorno, Fig. 4, sono state applicate le condizioni di simmetria totale per quanto concerne i bordi di taglio della lastra (quindi $x=0$ ed $y=0$) mentre per tenere conto della presenza degli dispositivi di fissaggio del provino sul supporto, Fig. 2, sono stati introdotti dei vincoli rispetto alla coordinata z nella zona più esterna della lastra, Fig. 4. I valori dei carichi che sono stati applicati sono sempre stati desunti dalle prove sperimentali ed il carico è stato applicato ai nodi dell'elemento al centro della piastra su un diametro ideale di contatto pari a 3mm. Infine si fa notare che per la soluzione dei modelli un aspetto particolarmente critico è stata la scelta del numero di passi per la soluzione del problema non lineare. In particolare nel caso dei modelli che contengono la lastra porcellanata si è resa necessaria una soluzione in due fasi: nella prima

fase è stata effettuato un calcolo con carico termico uniforme (salto termico imposto di 400°C), nella seconda fase mantenendo attivo il salto termico è stato introdotto il carico meccanico. Le due fasi sono state tra di loro congiunte mediante una operazione nota come “restart”. Inoltre, sempre nel caso delle lamiere porcellanate si è reso necessario introdurre l’operazione di “death” degli elementi del rivestimento che superano il criterio della massima tensione al fine di rendere conto della frattura del medesimo.

I modelli che sono stati sviluppati sono i seguenti:

1. lastra in solo acciaio avente spessore uniforme di 0.8 mm (SIM-1)
2. lastra in solo acciaio avente spessore uniforme di 1.2 mm (SIM-2)
3. lastra in acciaio, di spessore 0.8 mm, rivestita con smalto porcellanato su entrambi i lati con spessore di 0.2 mm di rivestimento per lato (ottenendo quindi uno spessore totale del modello pari a 1.2 mm); per tale configurazione sono state distinte due condizioni di analisi:
 - 3.a. lastra rivestita priva di tensioni residue (SIM-3);
 - 3.b. lastra rivestita soggetta a tensioni residue (SIM-4).

Lo studio numerico è stato sviluppato considerando i quattro modelli ora citati al fine di avere una base dati confrontabile con le prove sperimentali; inoltre l’ultimo modello, la lastra porcellanata nella quale agiscono le tensioni residue, è stato utilizzato per svolgere una sperimentazione numerica aggiuntiva. Tale sperimentazione ha avuto lo scopo di verificare l’influenza della tensione residua sulla risposta della lamiera porcellanata e a tal fine sono stati considerati differenti valori del coefficiente di dilatazione termica: $\alpha \in [1.00; 1.05; 1.10; 1.15; 1.20; 1.25] \times 10^{-5} \text{ } ^\circ\text{C}^{-1}$.

3. RISULTATI E DISCUSSIONE

La presentazione dei risultati e la loro discussione viene articolata in due fasi: preliminarmente vengono presentati ed analizzati i risultati relativi alle prove sperimentali, nel secondo paragrafo vengono presentati e commentati i risultati delle simulazioni numeriche.

Risultati e discussione delle prove sperimentali

Nella Figura 5 sono riportati quattro esempi di diagrammi delle superfici delle lamiere smaltate e non che si sono formate a seguito dell’impatto. I diagrammi in Figura 5 sono stati ottenuti mediante scansione laser delle superfici impattate.

Le prove effettuate sulle lamiere non smaltate hanno messo in evidenza l’insorgere di una instabilità localizzata nei bordi liberi della lastra. Da tali diagrammi si osserva che l’instabilità non insorge al primo di livello di energia scelto, corrispondente ad una energia di 3J, mentre si manifesta a seguito di impatto da una altezza pari a $H=0.50\text{m}$ con una energia di 6J.

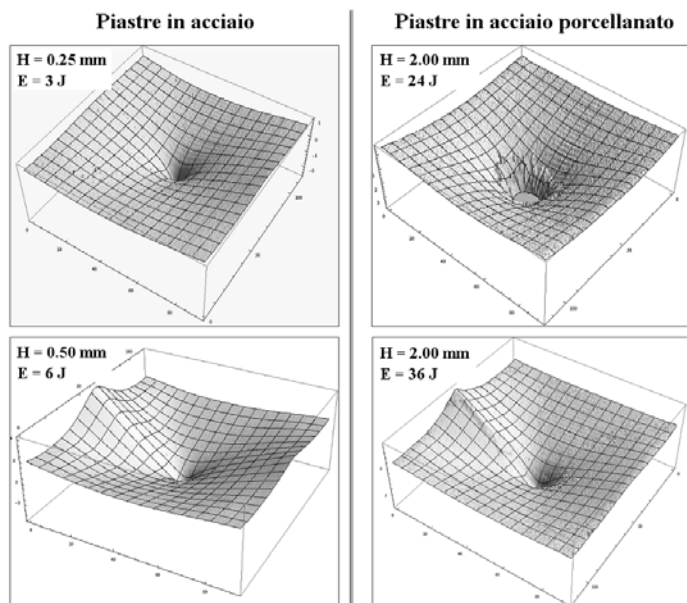


Figura 5. grafici delle superfici deformate dopo impatto ottenute mediante scansione laser

Una attenta analisi della letteratura ha permesso agli autori di osservare che tale fenomeno è stato messo in evidenza nel 1954 dal Belluzzi [10] il quale ne fornisce una buona descrizione fenomenologica ma non una descrizione analitica esatta. Inoltre le analisi e le relazioni semiempiriche che egli propone sono riferite al caso di instabilità di lastre circolari mentre viene solo accennato che un simile fenomeno si verifica anche nel caso di lastre rettangolari.

Dal punto di vista fisico, come osserva il Belluzzi, il fenomeno di instabilità che nasce nella lamiera soggetta alla condizione di carico e vincolo oggetto di studio è dovuto ad un complesso meccanismo di trasmissione delle azioni interne alla piastra che porta alla nascita di forze di compressione agenti nel piano della piastra medesima e trasversali alle bisettrici del rettangolo.

Infine si osserva che in base alle geometrie post impatto che hanno assunto le lamiere che presentano instabilità si può ipotizzare che il processo di impatto ha portato a lavorare il materiale oltre il limite elastico.

Nel caso delle lamiere porcellanate l'insorgere della instabilità, come evidenziato in Figura 5 e Figura 6, avviene ad un livello molto superiore in termini di altezza di caduta dell'impattatore e quindi di energia di impatto.

In particolare dall'analisi delle fotografie in Fig. 6 si osserva come in tutti i casi riportati il rivestimento ceramico vetroso subisca un importante danneggiamento nella zona centrale pur tuttavia rimanendo sostanzialmente integro nelle zone ove la lamiera non porcellanata ha esibito l'instabilità.

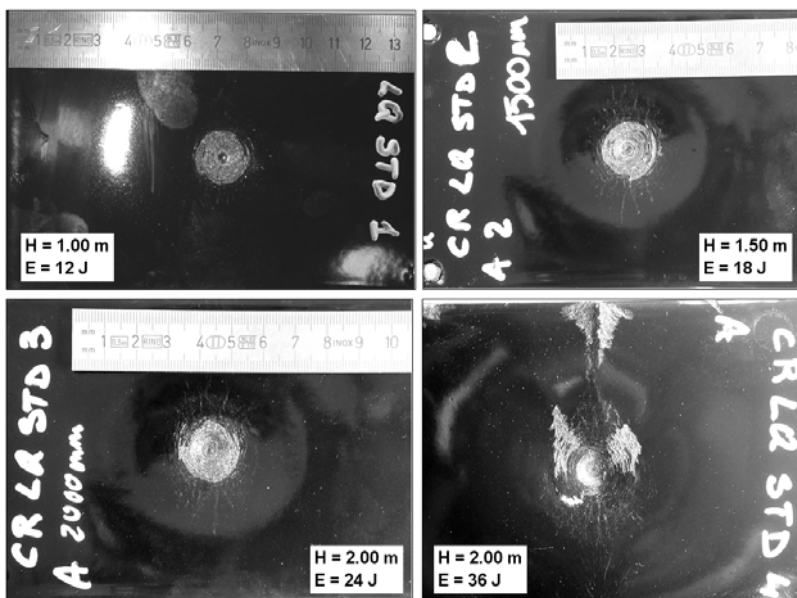


Figura 6. immagini fotografiche relative a quattro esempi di lamiere porcellanate impattate; nel caso di impatto con energia pari a 36J si verifica l'insorgere della instabilità.

Si osserva inoltre che nei casi di $E=12J$, $E=18J$ ed $E=24J$ il danneggiamento oltre ad interessare la zona centrale con il distacco del rivestimento si estende anche alle zone circostanti sotto forma di cricche che però non causano il distacco del rivestimento stesso.

Soltanto il maggiore livello di energia di impatto rivela la perdita del rivestimento nelle zone ove si è manifestata l'instabilità. Inoltre confrontando le immagini con $E=18J$ ed $E=24J$ con $E=36J$ si osserva come nelle prime siano già presenti delle linee di frattura che preludono alla perdita per distacco dello smalto nella direzione lungo la quale l'ultima immagine ($E=36J$) mostra l'insorgere della instabilità.

Le prove sperimentali relative agli impatti hanno fornito informazioni relative alla risposta meccanica rispetto all'evoluzione temporale del carico ed alla risposta in termini di carico-spostamento. Inoltre come parametro di interesse è stata anche considerata l'energia assorbita durante l'impatto.

Nei diagrammi di Figura 7 sono riportati gli esempi dei rilevamenti effettuati durante gli impatti sulle lamiere in acciaio. In particolare nel caso di $E=3J$ non si osservano evidenti discontinuità nell'andamento del carico (si nota per altro un andamento tipico a campana), diagrammi A e B, cosa che invece si presenta per le altezze di impatto maggiori. Dal diagramma di Figura 7-C si osserva che al crescere dell'energia di impatto la forza massima che si registra durante l'impatto medesimo ha un andamento che tende ad un valore asintotico orizzontale.

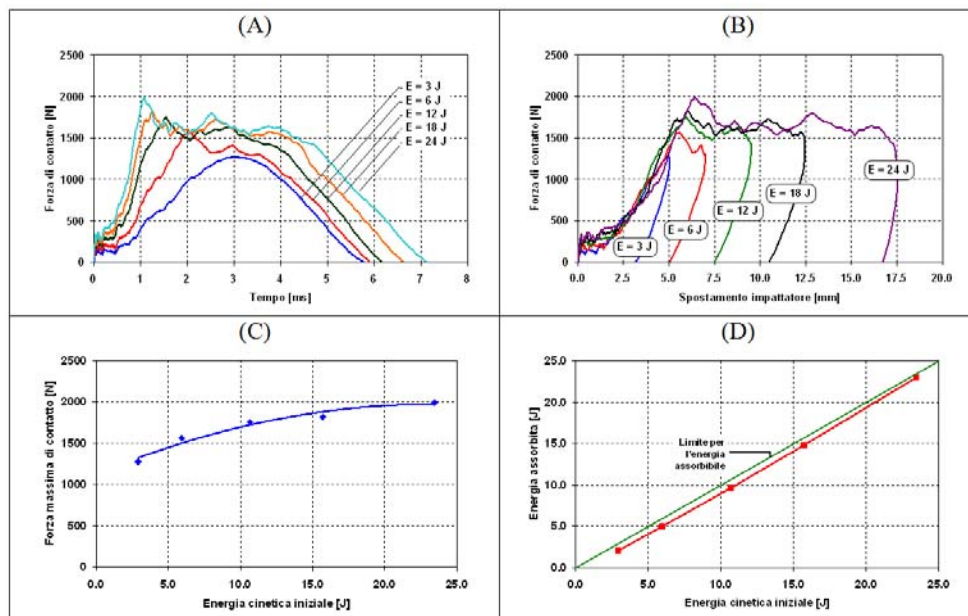


Figura 7. Diagrammi di impatto relativi alle *lamiere non smaltate*; (A) diagramma forza-tempo, (B) diagramma forza-spostamento, (C) diagramma forza di contatto massima-energia cinetica iniziale, (D) diagramma energia assorbita-energia cinetica iniziale

Per contro l'andamento dell'energia assorbita durante l'impatto rispetto all'energia cinetica iniziale, Figura 7-D, risulta poco scostato dal limite dell'energia massima assorbibile (pari all'energia iniziale, come evidenziato dalla retta in figura).

L'andamento della forza in funzione dell'energia cinetica iniziale, unitamente agli altri andamenti presenti nei diagrammi A e B, rivela il fatto che il materiale della piastra durante l'impatto viene portato a lavorare in campo plastico e che l'estensione dell'area plastica tende ad accrescersi in modo significativo.

I diagrammi in Figura 8 sono invece riferiti ai rilevamenti degli impatti sulle lamiere in acciaio porcellanato. Dai diagrammi in Figura 8-A e Figura 8-B, si osserva che soltanto nel caso della massima energia cinetica iniziale, $E=36J$, l'andamento della forza di contatto ha marcato crollo dopo il valore massimo. Tale crollo rivela l'innescò della instabilità che, come osservato in precedenza, è accompagnato dalla perdita di rivestimento nella zona ove l'instabilità si manifesta. In tal senso risulta interessante effettuare un confronto diretto tra gli andamenti della forza e dell'energia assorbita nel caso in cui si è verificata l'instabilità sia per la lamiera in acciaio che per quella in acciaio porcellanato. Tale confronto viene riportato in Figura 9.

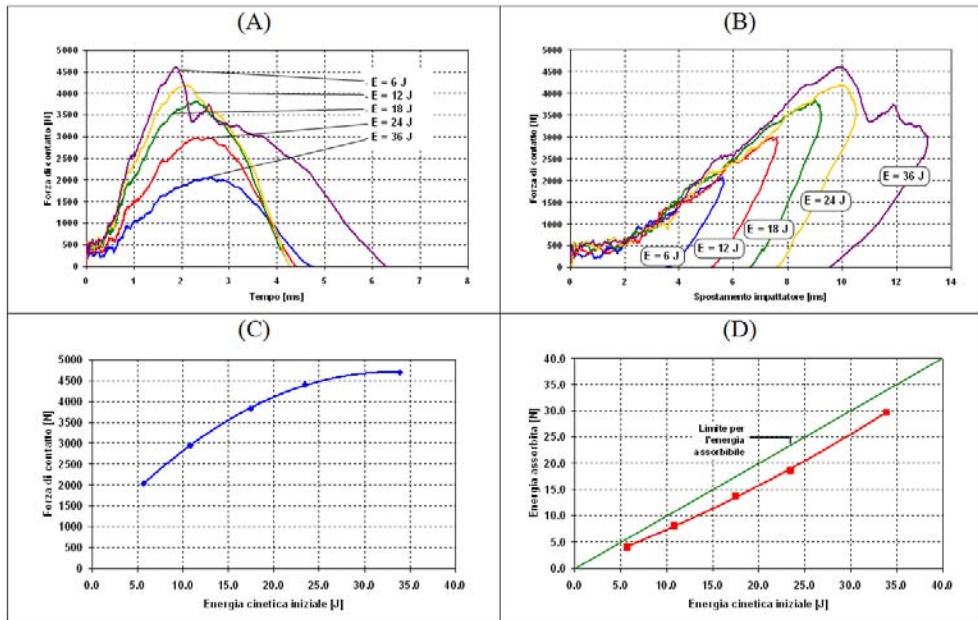


Figura 8. Diagrammi di impatto relativi alle *lamiere porcellanate*; (A) diagramma tempo-forza, (B) diagramma spostamento-forza, (C) diagramma energia cinetica iniziale-forza massima di contatto, (D) diagramma energia assorbita-energia cinetica iniziale

In particolare si osserva che sebbene le scale della forza e della energia assorbita siano nettamente differenti, Figura 9-A e Figura 9-C, passando a diagrammi su doppia scala, Figura 9-B e Figura 9-D, le similitudini morfologiche dei diagrammi vengono messe in evidenza.

Nel diagramma relativo alla forza di impatto si osserva in entrambi i casi, smaltato e non, la presenza di un brusco cedimento della forza dopo il picco di massimo a cui fa seguito un secondo picco di forza, di intensità inferiore rispetto al primo, dopo il quale la forza decresce progressivamente con un tipico andamento a campana. In modo del tutto analogo anche l'andamento delle energie assorbite, Figura 9-D, è molto simile. Si tiene a far notare che il processo fisico che è alla base del primo picco di forza può essere collegato allo snervamento del materiale mentre il secondo picco di forza, di intensità inferiore al primo, può essere dovuto ad un fenomeno di incrudimento e/o di irrigidimento associato alla instabilità locale. Per quanto attiene a questo fenomeno di incrudimento si rimanda all'analisi numerica per poter avere maggiori dettagli.

Infine nella Figura 10 vengono poste a confronto la forza massima e l'energia dissipata registrate durante l'impatto riferite all'energia iniziale.

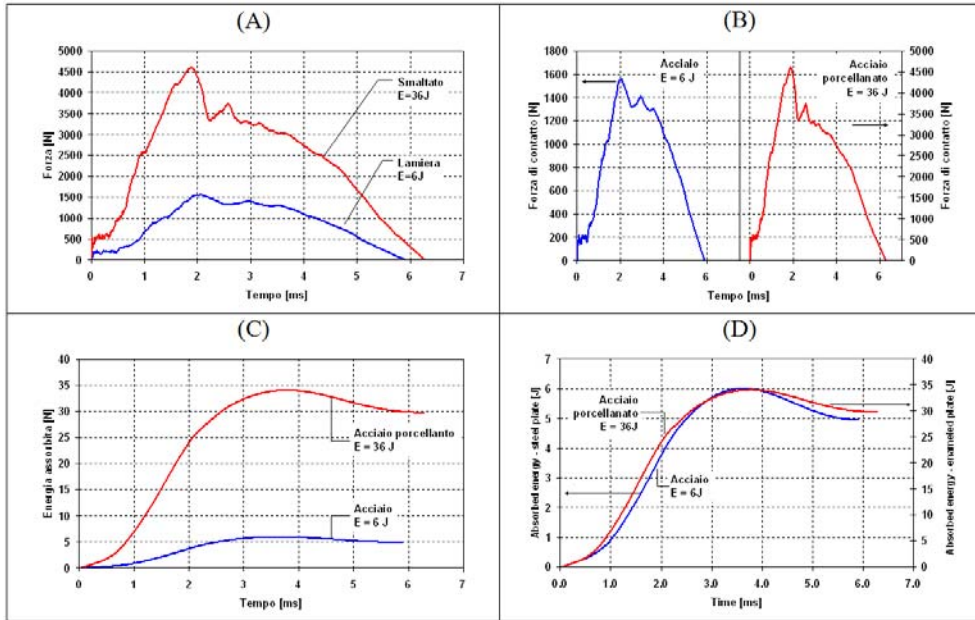


Figura 9. Confronto tra i diagrammi di forza-tempo ed energia-tempo relativi al caso di instabilità sia delle lamiere non smaltate che di quelle porcellanate

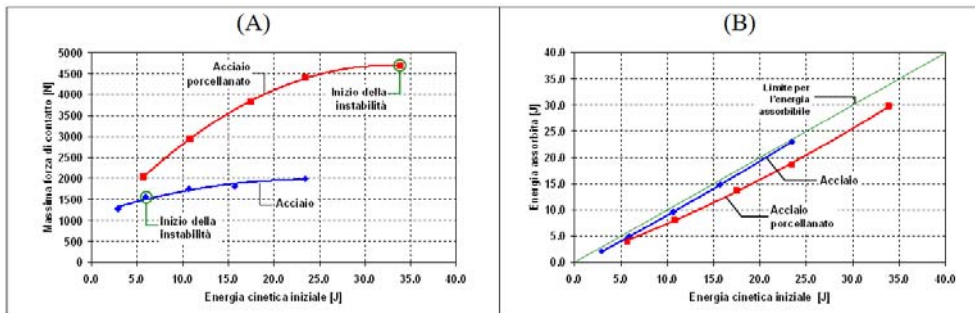


Figura 10. Confronto tra la forza massima e l'energia assorbita nel caso di lamiere smaltate e non per le diverse altezze di impatto.

Per quanto concerne la forza massima si osserva che essa risulta essere sempre maggiore nel caso delle lamiere porcellanate rispetto a quelle non rivestite, mentre per quanto attiene all'energia assorbita durante l'urto si osserva che essa è maggiore nel caso delle lamiere non smaltate rispetto a quelle porcellanate.

I due fatti ora citati e riassunti in Figura 10-A e Figura 10-B possono essere spiegati tenendo conto di due aspetti: le lamiere porcellanate hanno complessivamente uno spessore maggiore rispetto a quelle prive di rivestimento, inoltre, cosa che vedremo essere di una certa rilevanza, le lamiere porcellanate sono interessate da un sistema di tensioni residue. In particolare, come già accennato, nel caso delle lamiere porcellanate, il substrato in acciaio

è soggetto a tensioni di trazione mentre il rivestimento è soggetto a tensioni residue di compressione. Nel suo complesso, quindi, il sistema è soggetto ad un irrigidimento per pretensionamento interno che inevitabilmente lo porta a reagire all'impatto con una forza maggiore e nel contempo a dissipare una minore quantità di energia.

Risultati e discussione delle analisi numeriche

Le ragioni che rendono molto complessa la descrizione della instabilità osservata durante le prove sperimentali sono principalmente quattro: (i) l'instabilità è di tipo locale, ovvero non interessa tutta la lastra, (ii) essa deriva da un stato di sollecitazione interna alla lastra che viene indotto da un carico trasversale al piano stesso della lastra, (iii) si è in presenza di grandi spostamenti ed inoltre (iv) il tipo di vincolo cui è soggetta la lastra sui bordi è di tipo unilatero. La concomitanza di questi tre aspetti uniti alla non linearità del materiale che, nel nostro caso, è stata sperimentalmente verificata, fanno sì che l'approccio analitico sia ancora un problema aperto. Tuttavia, in coerenza con gli scopi del presente studio, si è deciso di effettuare una indagine numerica del fenomeno la più fedele possibile. In tal senso, come già descritto nel §.2, i modelli numerici che sono stati costruiti e risolti integrano i quattro aspetti ora citati.

Le analisi numeriche condotte sulla lamiera in acciaio, in condizioni di carico quasi statico, hanno permesso di mettere in evidenza l'insorgere e lo svilupparsi della instabilità. Nella Figura 11 vengono riportate due immagini relative alla stessa simulazione per una lastra in acciaio di spessore 0.8mm.

L'immagine in Figura 11-A-1 ed il diagramma ad esso associato, Figura 11-A-2, contengono rispettivamente la mappa dello spostamento e lo spostamento verticale agli apici della lastra in corrispondenza del sub-step dopo il quale inizia a manifestarsi l'instabilità. Mentre l'immagine ed il diagramma in Figura 11-B-1 e Figura 11-B-2 rappresentano invece, rispettivamente, la mappa dello spostamento e gli spostamenti verticali agli apici della lastra in corrispondenza dell'ultimo sub-step di calcolo. Nel diagramma dello spostamento U_z_1 in Figura 11-B-1 si osserva marcato cambiamento della pendenza che indica l'insorgere della instabilità della lastra.

Per poter analizzare l'instabilità e quindi identificare il carico in corrispondenza del quale essa viene indotta, si è scelto di procedere mediante interpolazione dei due tratti relativi al diagramma dello spostamento U_z_1 in funzione della forza, Figura 12. I tratti che vengono interpolati si trovano rispettivamente prima e dopo la variazione di pendenza del diagramma spostamento-forza.

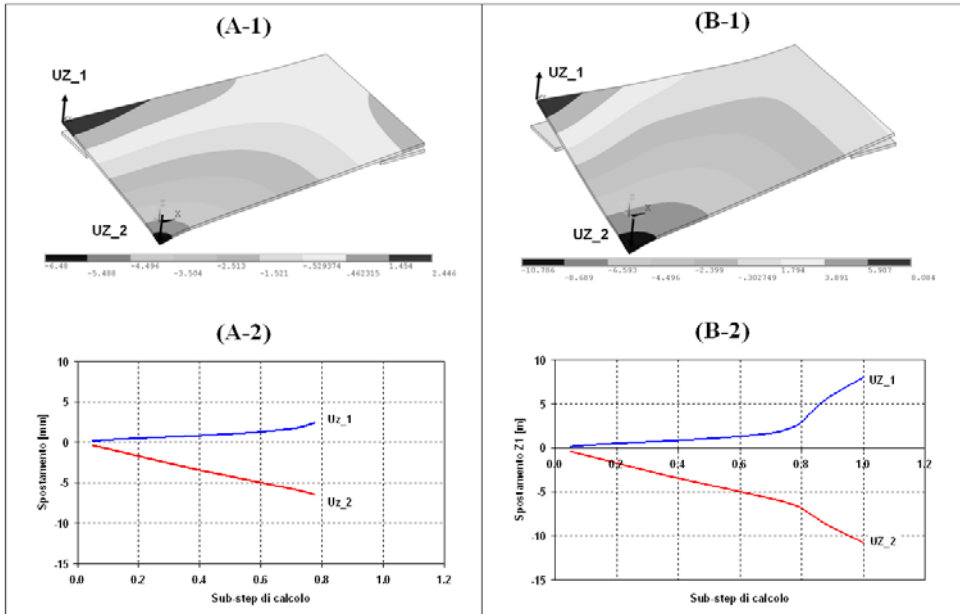


Figura 11. Immagini relative alla simulazione per la lastra in acciaio da 0.8mm di spessore corrispondenti a due sub-step della soluzione; (A) sub-step intermedio, (B) sub-step finale

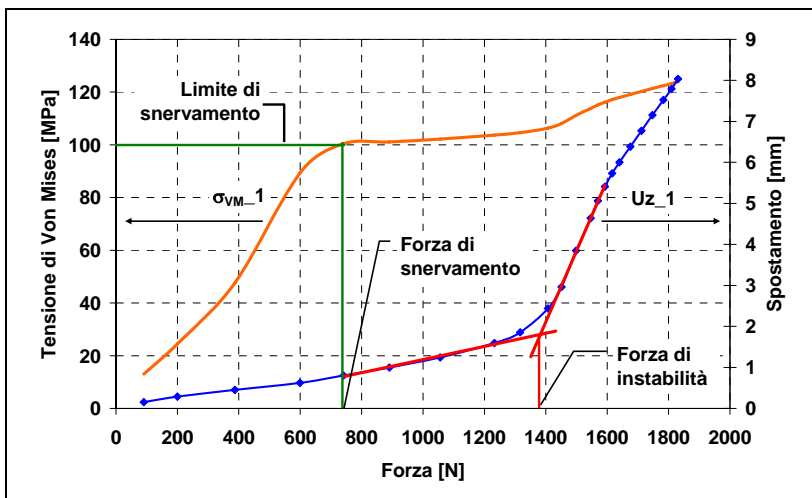


Figura 12. (A) andamento degli spostamenti di apice rispetto alla forza applicata; (B) andamento dello spostamento di apice UZ_1 e della tensione di Von Mises

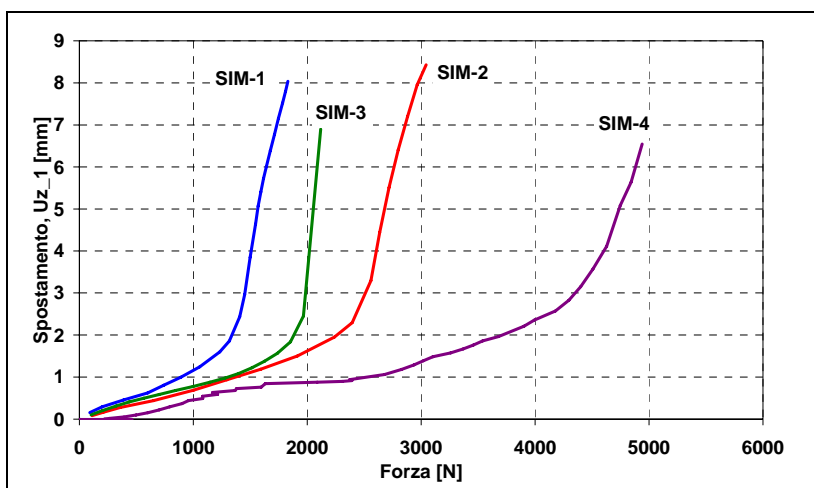


Figura 13. Andamenti dello spostamento Uz_1 per i quattro modelli in funzione del carico

Tabella3: valori relativi ai carichi critici ottenuti mediante le simulazioni numeriche

ID	Spessore Lamiera [mm]	Smalto	Max tensione residua [MPa]		F_{cr} [N]
			Acciaio	Smalto	
SIM-1	0.8	NO	\	\	1361
SIM-2	1.2	NO	\	\	2390
SIM-3	0.8	SI	0	0	1909
SIM-4	0.8	SI	20.2	-37.8	4400

Nella Figura 13 vengono riportati i diagrammi relativi agli andamenti dello spostamento Uz_1 in funzione della forza applicata nei quattro casi previsti in §2, mentre nella Tabella 3 vengono riassunti i valori delle forze che inducono l'instabilità (F_{cr}).

Una prima osservazione relativa ai valori di carico critico ottenuti mediante le simulazioni numeriche riguarda i risultati di SIM-1 e SIM-4 che rappresentano i due casi oggetto anche di indagine sperimentale. Dalle prove sperimentali per le lamiere in acciaio si può dedurre che l'instabilità viene causata da un urto nel quale l'energia cinetica dell'impattatore è compresa tra 3J e 6J e che la forza massima di contatto che è associata alla condizione che causa l'instabilità è compresa tra 1275N e 1560N. Se si considerano i risultati numerici si osserva che la forza calcolata che causa l'instabilità è di 1360N, valore che risulta essere compreso nell'intervallo di forza massima rilevato sperimentalmente. Una analisi simile può essere fatta nel caso delle lamiere in acciaio porcellanato. In tal caso

infatti l'energia cinetica dell'impattatore che induce l'instabilità nella lastra è compresa tra 24J e 36J mentre la forza massima di contatto a cui è associata l'instabilità è compresa tra 4409N e 4694N. Dall'analisi numerica è emerso che il valore della forza che induce l'instabilità nella lastra in acciaio porcellanato è pari a 4400N, valore che risulta essere compreso nell'intervallo della forza massima di contatto individuata per via sperimentale. Da queste osservazioni si può quindi concludere che mediante le simulazioni numeriche è stato possibile individuare dei valori plausibili per la forza che causa l'insorgere della instabilità sia nel caso delle lastre in acciaio che in quelle in acciaio porcellanato.

Si osserva inoltre che:

- (i) all'aumentare dello spessore della lamiera il carico critico aumenta (confronto tra SIM-1 e SIM-2),
- (ii) la presenza del rivestimento privo di tensioni residue fa sì che il carico critico aumenti rispetto alla lamiera di spessore 0.8mm mentre si riduce rispetto al caso di lamiera avente spessore 1.2mm (confronto tra SIM-1, SIM-2 e SIM-3),
- (iii) la presenza delle tensioni residue che agiscono nella lamiera porcellanata tende ad aumentare il valore del carico critico rispetto al caso di lamiera avente spessore equivalente, 1.2mm, ma realizzata in acciaio (confronto tra SIM-2 e SIM-4).

Da queste osservazioni emerge l'influenza della presenza delle tensioni residue che agiscono nella lastra porcellanata sul carico critico che ne innesca l'instabilità. Questo aspetto è stato oggetto di una ulteriore indagine numerica che ha considerato differenti valori dei coefficienti di dilatazione termica degli smalti porcellanati, al fine di indurre differenti stati di tensione residua nella lamiera porcellanata. Nel diagramma di Figura 14 sono riassunti i risultati relativi a tali analisi numeriche.

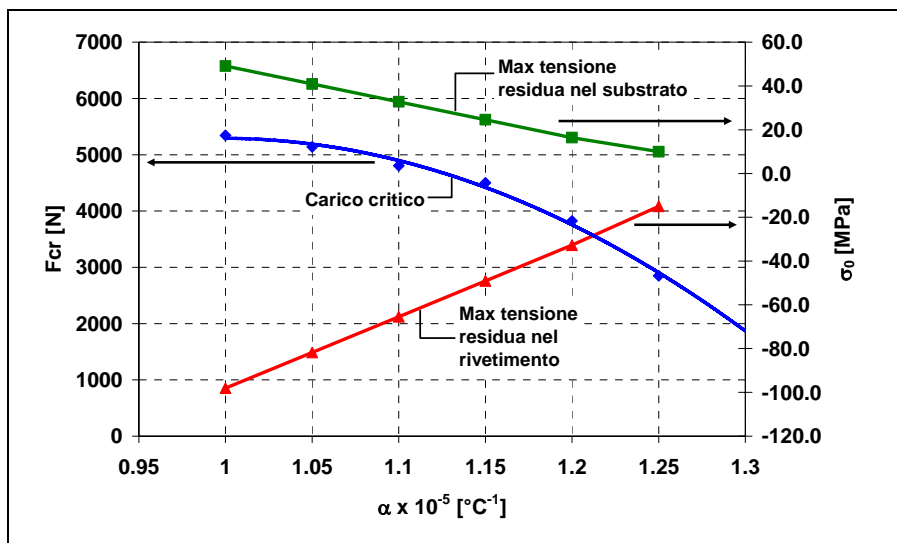


Figura 14. andamento del carico critico al variare del coefficiente di dilatazione termica del rivestimento e relativi andamenti delle tensioni residue nel substrato e nel rivestimento

Dai diagrammi in Figura 14 si osserva che al crescere del valore del coefficiente di dilatazione termica le tensioni residue che agiscono nella lamiera porcellanata calano (il rivestimento tende a comportarsi dal punto di vista della dilatazione termica come il substrato) e contestualmente calano anche i valori relativi al carico critico.

4. CONCLUSIONI

Lo studio sperimentale del comportamento all'impatto a bassa velocità di lamiere, di spessore 0.8mm, in acciaio a bassissimo tenore di carbonio, smaltate e non, ha permesso di mettere in evidenza l'insorgere di una instabilità locale nelle lamiere. L'instabilità compare sul lato più lungo dei provini rettangolari considerati e si manifesta in modo apprezzabile solo quando l'energia di impatto supera un valore di soglia critica. I risultati sperimentali hanno permesso di osservare che l'energia critica per le lamiere porcellanate è circa 6 volte superiore a quella delle lamiere non rivestite. Le analisi numeriche agli elementi finiti di tale fenomeno sono state condotte considerando l'applicazione di un carico quasi statico al centro di quattro modelli di riferimento: lamiera in acciaio avente spessore pari a 0.8mm, lamiera in acciaio di spessore pari a 1.2mm, lamiera porcellanata (avente spessore del substrato pari a 0.8mm e spessore del rivestimento pari a 0.2mm per lato) nella quale non agiscono le tensioni residue e lamiera porcellanata nella quale agiscono le tensioni residue. In tutti i modelli per il materiale della lamiera è stata utilizzata una legge costitutiva bilineare mentre per lo smalto è stato assunto un comportamento lineare fragile con perdita di rigidità per gli elementi che superano la tensione di rottura. I risultati delle simulazioni hanno messo in evidenza l'insorgere della instabilità in corrispondenza di valori di carico in ottimo accordo con quelli ottenuti mediante le prove sperimentali. Inoltre i risultati delle simulazioni hanno permesso di valutare l'influenza delle tensioni residue sul valore del carico critico che provoca l'insorgere della instabilità. In particolare si è verificato che le tensioni residue innalzano il valore di tale carico critico. Infine è stato condotto uno studio quantitativo dell'influenza di differenti valori di tensione residua sul carico critico indotti da rivestimenti porcellanati aventi stesso modulo di Young ma differente coefficiente di dilatazione termica. Tale studio ha messo in evidenza una decrescita non lineare del carico critico al crescere del coefficiente di dilatazione lineare fino al limite della sua uguaglianza a quello dell'acciaio.

Ringraziamenti

Si ringrazia la ditta SMALTIFLEX S.p.A. per la preparazione dei campioni smaltati e l'assistenza nel corso delle prove di corrosione, la ditta COLOROBIA S.p.A. per le misure relative ai coefficienti di dilatazione termica. Infine un ringraziamento all'Ing. Gabriele Cuccol0069ni per l'esecuzione dei rilievi laser delle lamiere impattate.

BIBLIOGRAFIA

- [1] Andrews A. 1961, I. *Porcelain Enamel*, 2nd ed., the Garrard Press.
- [2] R. Ambu, A. Zucchelli, L. Rossetti, V. Dal Re, 2007, "Residual Stresses evaluation in Functionally Graded Composite Vitreous Enameled Steel Sheets", Proceedings of ETDCM8 - 8th Seminar on Experimental Techniques and Design in Composite Materials 3-6 October 2007 - Sant'Elmo Beach Hotel - Castiadas - Costa Rei - Sardinia (Italy).

- [3] R. Poletti, A. Zucchelli, A. Chelli, 2005, “Experimental Investigation on Corrosion Resistance of Porcelain Enamel Composite Coating for Regenerative Air Heaters Parts”, in Proceedings of Danubia Adria Symposium 2005 – Monticelli Terme, Parma (Italy).
- [4] A. Chelli, R. Poletti, L.Pignatti, A. Zucchelli, L. Rossetti, V. Dal Re, S. Curioni, D. Prandstraller, S. Tiberi Vipraio, G. Palombarini, 2006, “Studio delle Proprietà Meccaniche e Tribologiche di Lamiere in Acciaio rivestite mediante Smalti Porcellanati”, Smalto Porcellanato, n.3, pp. 70-92.
- [5] A. Chelli, R. Poletti, L. Pignatti, F. Bruscoli, G. Pasqualetti, F. Bruni, A. Zucchelli, L. Rossetti, F. Lotti, G. Minak, V. Dal Re, S. Curioni, 2008, “Composite enameled steel elements for air preheaters and gas-gas heaters: an integrated approach from sheet forming and enamelling to basket assembly”, Proceedings of XXI° International Congress on Porcelain Enamel, pp. 130-158, 18-22 maggio, Shanghai-Cina
- [6] ANSALDO, 1987, “Enamel coating for heating elements”, internal report.
- [7] M.T. Kim, S.Y. Chang, O.Y. Oh, J.B. Won, H.W. Park, 2007, “Failure analysis of enamel-coated carbon steel heating elements of gas-gas heater for flux gas desulfurization system”, Engineering Failure Analysis 14, 686–693.
- [8] L. Rossetti, A. Zucchelli, V. Dal Re, 2007, “Caratterizzazione a Flessione e a Trazione di Rivestimenti Ceramico-Vetrosi applicati a Substrati Metallici”, Proceedings of XXXVI Convegno Nazionale AIAS – 4-8 Settembre, Università degli Studi di Napoli Federico II – Seconda Università degli Studi di Napoli.
- [9] ASTM D 7136/D 7136M, 2007, “Standard Test Method for Measuring the Damage Resistance of a Fiber-Reinforced Polymer Matrix Composite to a Drop-Weight Impact Event”, November.
- [10] O. Belluzzi, 1954, “Una singolare instabilità delle lastre sottili appoggiate al contorno (a strange instability of thin plates simply supported)”, Boll. Genio Civile, fasc. 1.
- [11] ANSYS user’s manual, version v.11.

METODI DI PROGETTAZIONE INTEGRATA E SIMULAZIONE DI SOTTOGRUPPI DI MACCHINE ALTERNATIVE

Angelo Oreste Andrisano, Giovanni Barbanti, Marcello Pellicciari, Francesco Leali,
Marco Faretra, Francesco Gherardini, Fabio Pini, Alberto Vergnano
Dipartimento di Ingegneria Meccanica e Civile
Università di Modena e Reggio Emilia
E-mail: angelooreste.andrisano@unimore.it

Sommario. *I moderni criteri di progettazione dei sistemi meccanici, spesso finalizzati a fronteggiare le elevate prestazioni delle macchine e la drastica riduzione dei costi e dei tempi di consegna, richiedono di considerare ed ottimizzare in tempi ridotti numerosi parametri funzionali all'interno di scenari operativi eterogenei, in cui gli aspetti strutturali, cinematici e dinamici non possono essere valutati indipendentemente. I metodi tradizionali di progettazione spesso affrontano questi aspetti in modo necessariamente approssimato e richiedono onerosi test su prototipi reali per la verifica e la messa a punto del sistema. La prototipazione virtuale utilizza diversi strumenti CAE (Computer Aided Engineering) per realizzare modelli di sistemi meccanici reali che tuttavia vengono ancora prevalentemente utilizzati come strumenti di verifica di sistemi già definiti con tecniche tradizionali. Nel presente lavoro si discute un metodo di progettazione che, pur fondandosi su solide basi di conoscenza tradizionali, permette di identificare con certezza i principali parametri progettuali su cui intervenire ed utilizza estesamente la prototipazione virtuale come strumento di progettazione. Si presenta inoltre la sua applicazione nella progettazione di un sistema per la sperimentazione di una macchina alternativa.*

Parole chiave: *metodi di progettazione, prototipazione virtuale, macchine alternative*

1. INTRODUZIONE

La competizione globale richiede di progettare sistemi meccanici dalle prestazioni sempre più elevate nonché il rispetto di grande affidabilità e di ridotti tempi di fermo macchina per interventi manutentivi. Parallelamente impone severi obiettivi di riduzione dei costi e stringenti vincoli di tempo, finalizzati a favorire il successo dei prodotti attraverso la capacità di fornire risposte concrete anticipate rispetto alla concorrenza, senza incorrere in penali legate al mancato rispetto dei tempi contrattuali di consegna che possano intaccare la redditività dei progetti.

Il processo di progettazione è quindi spinto verso la ricerca di architetture topologiche condivise e standardizzate, dalle quali derivare soluzioni caratterizzate da un elevato livello di personalizzazione delle funzioni e delle configurazioni degli organi accessori.

Tale scenario operativo impone la necessità di adottare metodi di progettazione integrata efficaci e robusti [1], in grado di evidenziare sin dalle prime fasi potenziali criticità e fonti di disturbo o depauperamento delle prestazioni/qualità finale del prodotto (Taguchi), nonché stimolare e supportare l'adozione di soluzioni e varianti progettuali alternative.

È ben noto, infatti, come, in prima approssimazione, i 2/3 dei costi da sostenere vengono implicitamente deliberati nelle prime fasi di sviluppo concettuale e di progettazione esecutiva (Fig. 1). La struttura di costo di un prodotto, quindi, viene prevalentemente determinata nelle prime fasi di sviluppo ed ingegnerizzazione mentre le spese vengono generalmente sostenute in fasi successive del ciclo di vita. Ne deriva che più si avanza nel processo di sviluppo, minori risultano essere i margini per l'introduzione di modifiche, dal momento che queste comportano costi sempre più elevati.

L'impiego di strumenti avanzati di simulazione e prototipazione virtuale [2], [3], [4], e la loro integrazione all'interno del ciclo di sviluppo prodotto divengono quindi strategici ma devono essere regolati da specifiche metodologie di definizione e di gestione del processo progettuale, fondamentali per raggiungere gli ambiziosi obiettivi precedentemente ricordati.

Allo stato dell'arte accade che lo sviluppo delle fasi di design concettuale e di primo approccio sia basato soprattutto sull'analisi di parametri progettuali identificati attraverso conoscenze empiriche e successivamente sviluppati mediante pratiche di dimensionamento [5], realizzate attraverso calcoli analitici basati su modelli semplificati ed ipotesi approssimative. Solamente nelle fasi finali dello sviluppo del progetto si procede alla verifica e all'ottimizzazione dei componenti attraverso l'uso intensivo di raffinati strumenti di simulazione, impiegati con un approccio di tipo "trial and error" [6].

Il recupero della conoscenza viene spesso fatto implicitamente, e non sempre del tutto consapevolmente, adottando standardizzazioni generiche e partendo da componenti e sottogruppi già realizzati in altri prodotti, perdendo, di fatto, la possibilità di identificare realmente i parametri su cui poter agire per adattare coerentemente le soluzioni preesistenti ai casi reali ed ottimizzarne il comportamento.

Inoltre, le ottimizzazioni finali, sviluppate con strumenti di simulazione fortemente specialistici e focalizzati su singoli aspetti comportamentali, impongono forti oneri computazionali e lunghi tempi di calcolo che ne limitano ulteriormente il campo applicativo, intervenendo su configurazioni di prodotto e soluzioni progettuali ormai definite e ad un elevato livello di avanzamento.

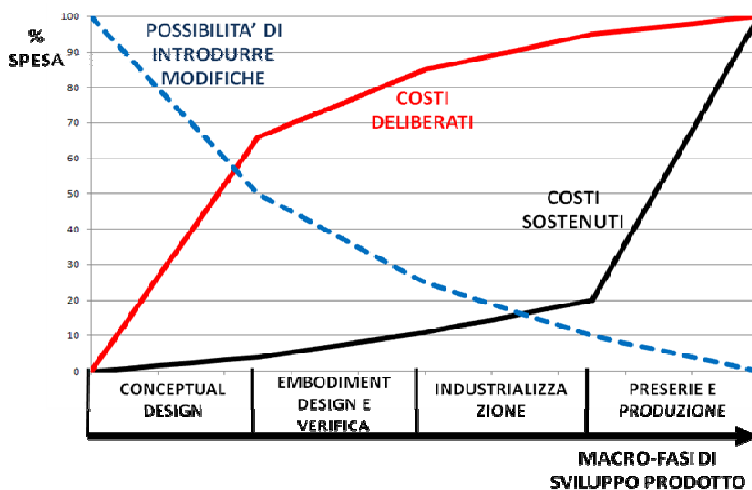


Figura 1. Costi deliberati e sostenuti nelle macro – fasi di sviluppo prodotto.

Proprio per la loro specificità, tali simulazioni, per quanto raffinate, non possono comunque evitare la realizzazione di prototipi fisici in scala su cui validare l'architettura progettuale e verificare le scelte adottate in relazione ai parametri funzionali inizialmente definiti, confermando, oltretutto una radicata prassi contrattuale.

La validazione progettuale realizzata attraverso la sola sperimentazione su prototipi fisici in scala, anche quando condotta secondo i principi applicativi del Design of Experiments (DOE), costituisce, spesso, un'ulteriore complicazione e induce ritardi legati ad inevitabili rotture di componenti, tali da imporre drastiche revisioni progettuali in fasi estremamente avanzate del processo di sviluppo. Infine, la stessa realizzazione in scala del prototipo necessita della decodifica dei fenomeni riscontrati e della trasposizione dei risultati nella scala del prodotto finale, portando talvolta a errori e malinterpretazioni anche clamorose. Solo l'accurata progettazione del processo di sperimentazione, l'identificazione dei parametri funzionali e delle relative correlazioni, la corretta scelta dei trasduttori necessari al monitoraggio delle grandezze in gioco ed il relativo posizionamento sul prototipo garantiscono l'affidabilità dei risultati ma rappresentano un compito oneroso e delicato e comportano lunghe ed accurate analisi dei dati sperimentali, occupando ulteriore tempo prezioso. Appare quindi evidente come la stessa progettazione del prototipo fisico in scala sia un'attività strettamente connessa al ciclo di vita del prodotto e necessiti non solo di approcci rigorosi ed efficaci ma di un'attenta analisi preliminare mirata a governarne lo sviluppo, in modo che i dati ricavati dalla sperimentazione possano essere realmente impiegati per la messa a punto dei parametri di simulazione [7][8][9].

Il processo progettuale risulta, in definitiva, talmente condizionato da vincoli di tempo e costo e così dipendente dagli stati di avanzamento in cui si identificano errori, potenziali alternative e modifiche migliorative che spesso i team di sviluppo decidono di non abbandonare soluzioni intrinsecamente non ottime pur di rispettare le scadenze fissate [10].

2. PROGETTAZIONE DI MECCANISMI DI MACCHINE ALTERNATIVE

I meccanismi che compongono le macchine alternative richiedono spesso processi di selezione e dimensionamento complessi, necessari a soddisfare specifiche di progetto solitamente molto spinte in termini di prestazioni ed affidabilità. Questi meccanismi operano in condizioni ambientali gravose e sono sollecitati con carichi molto elevati, derivanti dalla necessità di ottenere elevate potenze specifiche per dati regimi di funzionamento.

La progettazione tradizionale sfrutta formule analitiche [11], [12] per dimensionare i componenti dei meccanismi. Sebbene reso sicuro ed efficace da una solida tradizione ingegneristica, questo approccio introduce, inevitabilmente, diverse approssimazioni [6].

Innanzitutto il dimensionamento dei componenti si limita a configurazioni geometriche specifiche (punti morti inferiore e superiore, quadratura, etc.) e a particolari condizioni di funzionamento, trascurando, per limitare i tempi, aspetti importanti derivanti da condizioni di funzionamento particolari, come carichi secondari, forzanti e risonanze ai bassi regimi, nei transitori, o durante anomalie.

Le forze d'inerzia prese in considerazione nel processo di progettazione tradizionale spesso non sono quelle reali, ma vengono stimate solo in prima approssimazione, specialmente nel caso di geometrie molto complesse. La distribuzione delle masse ed i valori di accelerazione vengono valutati solo per stati di funzionamento notevoli del sistema. D'altra parte, dal momento che raramente sono disponibili le curve di carico complete agenti sui componenti, risulterebbe molto difficile valutare la compresenza di diversi carichi al di fuori di tali particolari condizioni di funzionamento.

Le approssimazioni che ne derivano sono frutto dell'esperienza dei progettisti e guardano al problema da un punto di vista conservativo, portando in certi casi (es. dimensionamento per la resistenza a fatica) al sovradimensionamento delle strutture rispetto a quanto necessario. I problemi strutturali e dinamici vengono separati sia per quanto riguarda le forze d'inerzia e l'interazione tra i componenti, sia per quanto riguarda i problemi vibrazionali. Gli aspetti dinamici più complessi sono valutati accuratamente solo per i componenti maggiormente studiati dalla letteratura (es. bielle o alberi motore [5]), mentre risultano difficili da valutare le interazioni dinamiche tra gli stessi componenti.

La complessità dei fenomeni obbliga spesso ad introdurre approssimazioni o addirittura a trascurare alcuni fattori in fase di progettazione, rimandando tutto alla sperimentazione e alle verifiche finali su prototipi reali, determinando i problemi di progettazione degli esperimenti e di analisi dei dati già citati in precedenza e comportando un inevitabile allungamento dei tempi di sviluppo.

Per riuscire a velocizzare le fasi di progettazione e sviluppo è innanzitutto necessario riuscire a formalizzare ed organizzare basi di conoscenza a cui poter accedere e riferirsi durante la progettazione di nuove soluzioni, nel dimensionamento di primo approccio e nel confronto e scelta di varianti di progetto. Disporre di database di regole che racchiudano passate esperienze e algoritmi comportamentali è vitale per velocizzare ed amplificare i criteri di scelta ragionata delle soluzioni e di impostazione del design; al tempo stesso, riuscire ad acquisire le preziose conoscenze empiriche e le esperienze di guasti e debolezze è un'opera complicata e onerosa, ma fondamentale.

3. METODO E STRUMENTI DI PROGETTAZIONE

L'approccio ideato ha due scopi primari: formalizzare basi di conoscenza, da implementare successivamente in veri e propri knowledge silos, e sfruttare le potenzialità dei tools di simulazione, sin nelle prime fasi di progettazione, su modelli virtuali concettualmente equivalenti per progettare proattivamente.

Per raggiungere il primo obiettivo si sono create una serie di librerie, basate essenzialmente su fogli di calcolo, in cui vengono formalizzate le regole di progettazione, le migliori prassi, gli algoritmi di dimensionamento e le teorie di calcolo più quotate che possono dinamicamente interagire e guidare le geometrie CAD parametrico variazionali. Anche per i casi più generali, quali ad esempio la modellazione di superfici di contatto, sono stati realizzati opportuni fogli di calcolo che implementano la teoria di Hertz e guidano per un ulteriore dimensionamento, un'analisi ad elementi finiti su modelli CAD fortemente semplificati e tali da rendere l'analisi comunque estremamente veloce per non interrompere il flusso creativo del progettista.

Per perseguire il secondo, ossia per favorire l'introduzione della prototipazione virtuale nelle prime fasi del processo di progettazione e ridurre drasticamente i tempi ed i costi dello sviluppo [13], si sono identificati strumenti di simulazione virtuale integrati all'interno di sistemi CAD 3D avanzati; tali sistemi permettono di analizzare e confrontare velocemente diverse soluzioni progettuali mantenendo un legame diretto con la geometria delle parti (Fig. 2).

Già da tempo sono disponibili sul mercato strumenti di simulazione evoluti e potenti, (es: [14], [15]), focalizzati però su singoli aspetti comportamentali che spesso nella realtà fisica sono mutuamente interagenti e che richiedono pertanto la creazione di rigide interfacce tra i gli applicativi. Per ovviare a tali inconvenienti sono stati scelti, invece, strumenti di tipo più generico, CAD 3D parametrici, che hanno permesso di variare le geometrie delle parti e, simultaneamente, le grandezze specifiche dei modelli matematici

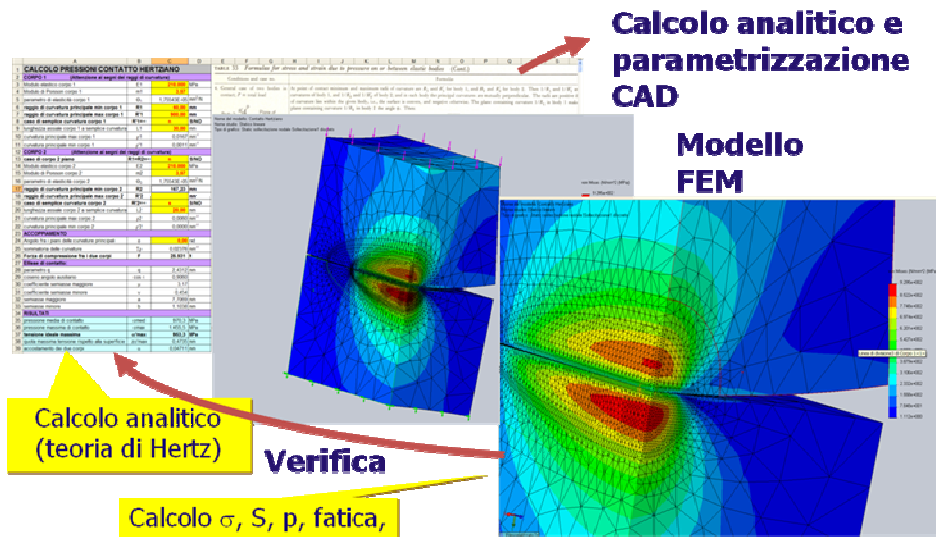


Figura 2. Dimensionamento di due profili a contatto Hertziano.

per simulare il funzionamento di diversi scenari operativi (“what if” scenarios), attraverso tecniche Design of Simulation (DOS, [16]).

I due strumenti più importanti per la prototipazione virtuale di meccanismi di macchine alternative sono risultati essere:

- L’analisi dinamica a corpi rigidi (MBD, Multi Body Dynamics, [2]).
- L’analisi strutturale ad elementi finiti (FEA, Finite Elements Analysis, [3]).

Com’è noto, l’analisi dinamica multibody (MBD) permette di determinare le risultanti delle interazioni tra i vari componenti e tra gli stessi componenti ed i vincoli esterni, le forze d’inerzia e tutto quanto necessario alla completa definizione della dinamica del sistema.

I software per l’analisi strutturale ad elementi finiti vengono utilizzati per determinare tensioni e deformazioni di ogni componente del sistema. Una fase molto delicata di queste simulazioni è la definizione dei vincoli. È importante che questi rappresentino effettivamente le reali interazioni tra i membri dei meccanismi. Sulle corrette aree di contatto si devono produrre distribuzioni di carico corrispondenti a quelle fisicamente scambiate tra i componenti reali. La simulazione delle reali interazioni sui vincoli interni ed esterni è di fondamentale importanza per la definizione di un prototipo virtuale rappresentativo di uno reale.

La scelta degli strumenti di simulazione da utilizzare viene necessariamente condizionata dalle loro potenzialità di integrazione. Infatti, se basati su un ambiente CAD 3D comune, i modelli virtuali possono essere rapidamente aggiornati ed utilizzati per diverse simulazioni. Nell’analisi di un prototipo virtuale le simulazioni MBD e FEA devono essere opportunamente integrate [4]. I carichi ricavati dalle simulazioni MBD devono essere utilizzati come input per i calcoli strutturali. Le simulazioni FEA devono invece fornire informazioni sulla ripartizione dei carichi in caso di iperstaticità, sulle deformazioni (specialmente nel caso in cui si considerino grandi deformazioni) e sul comportamento dinamico dei componenti (analisi modale). Queste devono poi essere

utilizzate in modelli a parametri concentrati oppure in simulazioni elastodinamiche [17] al fine di individuare i punti critici del sistema e le soluzioni tecniche da preferire, soprattutto in relazione al comportamento strutturale (resistenza e/o rigidità) e dinamico (traiettorie, accelerazioni, frequenze di risonanza, etc.).

In questo modo, assegnando opportunamente le leggi di moto si possono studiare velocemente le diverse condizioni di funzionamento (stazionarie o transitorie) e si possono ricavare ed integrare i dati relativi a tutte le condizioni operative ed a tutte le configurazioni del sistema. L'analisi dell'interazione tra i componenti può quindi tenere conto della reale trasmissione delle forze in tutte le condizioni operative [18].

Una volta identificati i principali parametri di interesse per l'ottimizzazione progettuale si possono definire modelli virtuali più raffinati che consentono di valutare in maniera attendibile le principali variabili del problema in tutti i punti più critici, anche quelli che sarebbero stati inaccessibili in un prototipo reale.

Sulla base di tale approccio sistemico [1] e delle scelte effettuate durante la definizione della catena di progettazione è stato quindi formalizzato un metodo suddiviso nei seguenti passi:

- a) Dimensionamento di massima del sistema mediante tecniche tradizionali, possibilmente formalizzate in fogli di calcolo elettronici.
- b) Disegno CAD 3D preliminare del sistema.
- c) Prototipazione virtuale del sistema:
 - determinazione dei carichi effettivi in tutte le condizioni operative
 - definizione dei vincoli interni ed esterni
 - analisi modale
 - calcoli di verifica strutturale
- d) Analisi dei risultati ed individuazione dei punti critici e delle necessarie modifiche.
- e) Ripetizione iterativa dei passi a-b-c-d per uno sviluppo progressivo del progetto fino a raggiungere un risultato completamente soddisfacente.

Le soluzioni palesemente inadatte vengono presto abbandonate senza dedicarvi ulteriori risorse mentre quelle più promettenti possono essere analizzate ad un maggiore livello di dettaglio. Vengono quindi apportate le necessarie modifiche ai punti critici del sistema, considerando sia il comportamento a regime che i transitori.

Seguendo tale metodo si possono eseguire, infine, analisi di sensibilità per valutare la robustezza del sistema nei confronti della variazione di parametri dimensionali (tolleranze) od operativi (dilatazioni, disturbi, etc.).

4. APPLICAZIONE E RISULTATI

Il metodo è stato validato attraverso la progettazione di un'attrezzatura di prova per componenti di macchine alternative. Lo scopo del dispositivo di prova è quello di sollecitare, con profili di carico assegnati, alcuni componenti di macchine alternative al fine di simulare particolari condizioni di funzionamento che si verificano sulla macchina reale. Per questa funzione è stato progettato il meccanismo di Fig. 3 il cui dimensionamento deve portare ad una forza di cerniera A da ottenersi in stretta tolleranza con le condizioni operative. Il componente che si muove di moto alterno viene caricato da un sistema elastico comandato da una camma, rigidamente fasata con la manovella, azionate a velocità angolare costante. Il rigoroso rispetto della legge di sollecitazione porta quindi a non poter prescindere, nei dimensionamenti, dagli aspetti connessi alla dinamica del sistema, e quindi agli effettivi valori dei carichi inerziali, all'elasticità di tutti i componenti ed agli aspetti deformativi.

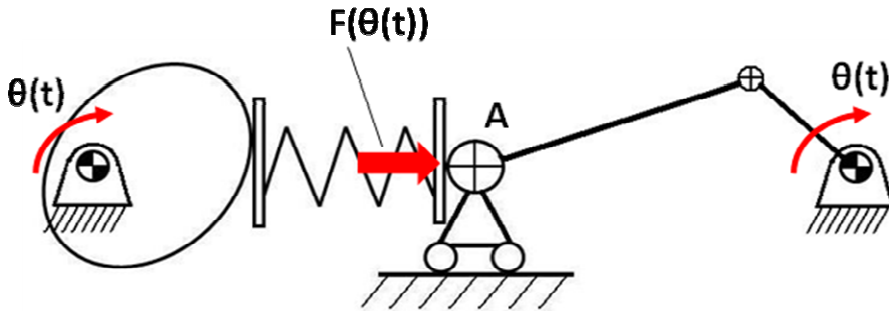


Figura 3. Schema cinematico dell'attrezzatura di prova.

La criticità del dimensionamento dei componenti, nonché la stretta tolleranza sui profili di forza applicati, richiedono il dimensionamento e l'ottimizzazione del sistema mediante prototipazione virtuale.

Il flusso di lavoro è quello descritto in Fig. 4. Il primo dimensionamento dei componenti viene effettuato mediante formule analitiche formalizzate in fogli di calcolo elettronici. In Fig. 5 è mostrato un esempio del dimensionamento della molla. Il modello matematico della molla viene generato automaticamente grazie ad un modello CAD 3D parametrico, direttamente collegato con il foglio di calcolo elettronico. La molla viene poi analizzata approfonditamente mediante FEA per studiare lo stato tensionale nelle sezioni. Le analisi mostrano che l'asse della molla si inclina durante il suo funzionamento determinando un sollevamento di alcune zone dei terminali. Le analisi MBD più dettagliate tengono conto dell'esatto punto di applicazione e della direzione effettiva della forza risultante esercitata dalla molla. Si calcolano inoltre le frequenze di risonanza della molla (vista come corpo continuo), per assicurarsi che queste non si trovino in prossimità delle frequenze di eccitazione del sistema.

Questo procedimento, seguito per tutti i componenti principali della macchina, porta al dimensionamento preliminare dei parametri del sistema, quali rigidità della molla, sezioni resistenti, masse in movimento, etc. Questi dati vengono utilizzati per la successiva impostazione di un modello MBD. Questo viene utilizzato per definire la forma dei profili a contatto e per ricavare dati su zone effettive di contatto, forze scambiate sulle superfici di contatto, deformazioni della molla, nonché effettivi valori delle grandezze dinamiche e cinematiche, coppie, velocità, accelerazioni, etc.

La molla e la sua struttura rappresentano un elemento critico per questo sistema e ne determinano le prestazioni in maniera significativa, sia per quanto riguarda la sua rigidità che per il suo comportamento dinamico (ha una massa elevata ed entrambi i suoi terminali sono in movimento durante il funzionamento). Essa può essere modellata solo in prima approssimazione come un sistema a massa e rigidità concentrate. Viene pertanto accuratamente analizzata con simulazioni elastodinamiche al fine di valutarne il comportamento dinamico e strutturale (Fig. 6).

Le simulazioni prevedono carichi e spostamenti imposti, derivanti dalle precedenti analisi MBD, per verificare anche eventuali fenomeni di risonanza.

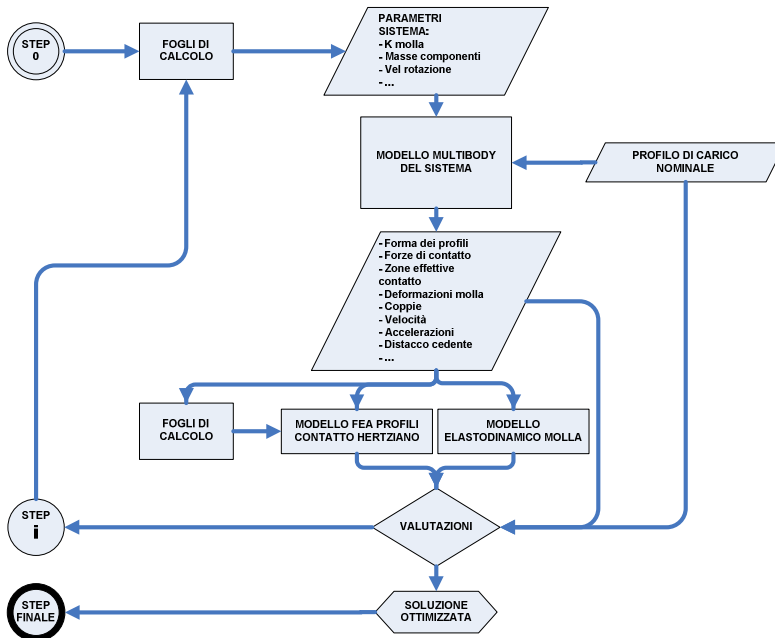


Figura 4. Flusso di lavoro per il dimensionamento del meccanismo a camma e molla.

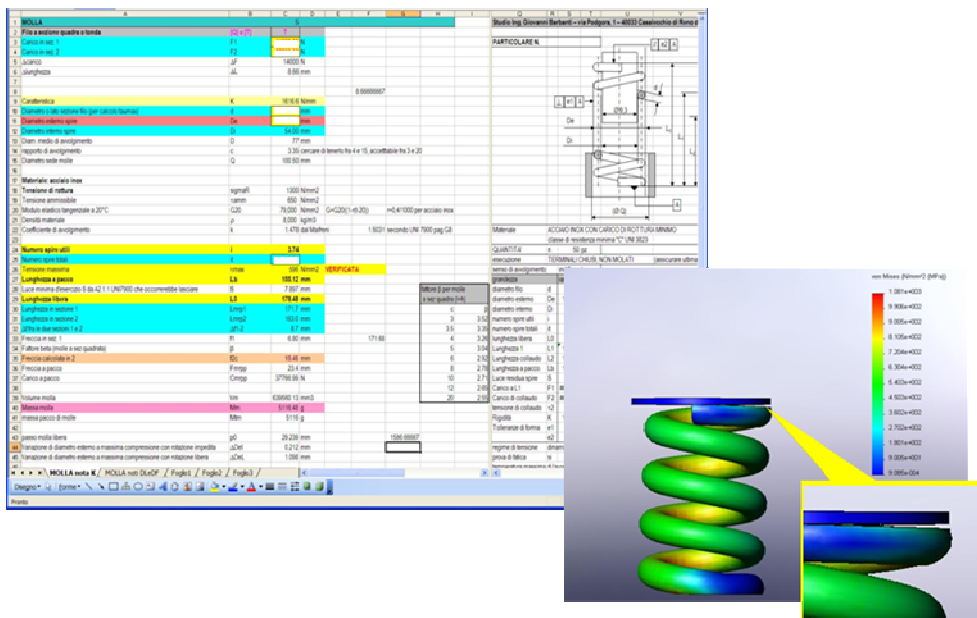


Figura 5. Dimensionamento della molla con foglio di calcolo ed analisi FEA.

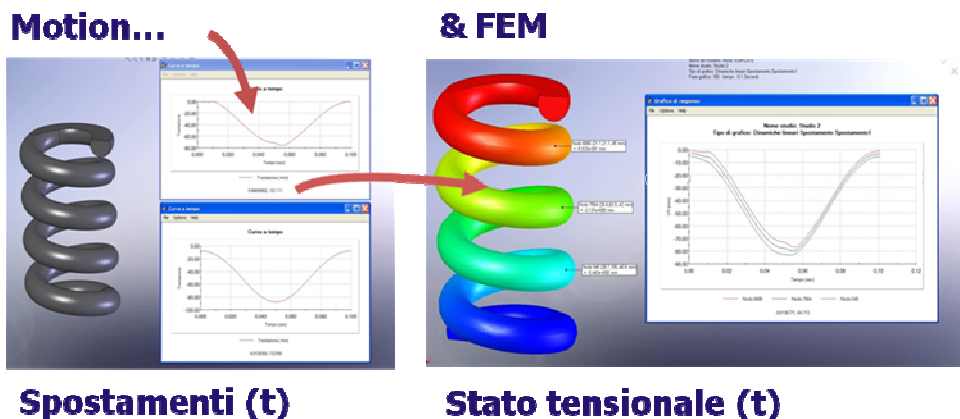


Figura 6. Analisi integrate dinamica e strutturale della molla.

Le analisi MBD e FEA hanno portato all'ottimizzazione dei parametri critici della molla. In particolare sono state valutate diverse rigidzze e, per iterazioni successive, si è ricercato un buon compromesso tra i seguenti aspetti:

- Facilità di messa a punto e sensibilità agli errori di montaggio
- Assenza di contatto tra i terminali e le spire
- Assenza di distacco del cedente dalla camma
- Picchi di pressione nel contatto Hertziano tra camma e cedente

Prima della prototipazione fisica dell'intero sistema è stata costruita la molla definitiva. Dopo la sperimentazione, le sue reali proprietà hanno permesso di rifinire gli ultimi dettagli dei componenti ad essa collegati grazie ad un'ultima simulazione virtuale.

5. DISCUSSIONE E SVILUPPI FUTURI

Il metodo proposto, basandosi sull'impiego di tecniche di calcolo numerico consolidate, consente di ridurre drasticamente il ricorso alle approssimazioni o semplificazioni spesso necessarie per una progettazione basata unicamente su metodi di calcolo analitico e di prendere in considerazione le reali condizioni di carico e di vincolo.

La corretta applicazione del metodo prevede l'impiego integrato degli strumenti di simulazione al fine di tener conto sul prototipo virtuale della contemporanea influenza di molti dei fenomeni che agiscono sul prototipo reale. L'applicazione del metodo ha portato ad una effettiva riduzione dei tempi di progettazione ed ottimizzazione prodotto. Inoltre è stato possibile valutare diverse soluzioni per i componenti collegati al sistema camma – molla al fine di arrivare ad una soluzione ottimale senza trascurare a priori altre soluzioni e senza alcun prototipo fisico preliminare. Si tratta di un metodo necessariamente iterativo poiché sfrutta tecniche di verifica, ma a differenza dei metodi tradizionali permette una visione completa del problema (anche se molto complesso) e un affinamento progressivo del progetto sulla base di una numerosa serie di analisi contestuali rese possibili dall'integrazione dei diversi strumenti di simulazione.

Il metodo presuppone la creazione di basi di conoscenza formalizzata che automatizzano i dimensionamenti di primo approccio: attraverso fogli di calcolo parametrici i cui risultati governano file CAD collegati si impostano dimensionamenti

preliminari che implementano conoscenze specifiche, riducendo tempi di sviluppo e aumentando la certezza operativa.

Il metodo permette poi di definire modelli che, una volta validati, possono essere facilmente adattati a nuovi casi di studio. Si creano quindi librerie di modelli validati ed immediatamente riutilizzabili. L'efficienza del metodo si incrementa progressivamente con l'uso. Queste basi di conoscenza, essendo agevolmente utilizzabili diventano patrimonio condivisibile nel caso di un processo di ingegneria concorrente.

La direzione indicata con questo metodo può facilmente integrare al proprio interno strumenti software per simulazioni fluidodinamiche, termodinamiche o dei sistemi di controllo. Questo potrà essere utile per valutare gli effetti, tutt'altro che trascurabili, delle interazioni tra fluido e struttura e le prestazioni accoppiate di sistemi integrati meccanici ed elettronici.

Ringraziamenti

Tale progetto è stato realizzato grazie al supporto delle strutture operative del laboratorio regionale dell'Emilia Romagna per l'Alta Meccanica INTERMECH-SIMECH.

Nota

L'incompletezza dei risultati presentati è conseguenza della clausola di riservatezza derivata dal coinvolgimento di una struttura industriale.

BIBLIOGRAFIA

- [1] Pahl, G., and Beitz, W., 2007. *Engineering design: a systematic approach*, 2nd ed.. Springer.
- [2] Boysal, A., and Rahnejat, H., 1997. "Torsional vibration analysis of a multi-body single cylinder internal combustion engine model". *Applied Mathematical Modelling*, **21**, August, pp. 481-493.
- [3] Reh, S., Beley, J.-D., Mukherjee, S., and Hui Khor, E., 2006. "Probabilistic finite element analysis using ANSYS". *Structural Safety*, **28**, pp. 17-43.
- [4] Lee, D.-C., and Han, C.-S., 2008 in press. "CAE (computer aided engineering) driven durability model verification for the automotive structure development". *Finite Elements in Analysis and Design*.
- [5] Ker Wilson, W., 1956. *Practical solutions of torsional vibration problems*, Vol. 3. Chapman & Hall Ltd, London.
- [6] Roy, R., Hinduja, S., and Teti, R., 2008. "Recent advances in engineering design optimization: Challenges and future trends". *CIRP Annals – Manufacturing Technology*, **57**, pp. 697-715.
- [7] Chu, C.-H., and Trethewey, M. W., 1998, "Rapid structural design change evaluation with an experiment based FEM". *Journal of Sound and Vibration*, **211**, pp. 335-353.
- [8] Kelton, W. D., 2000. "Experimental design for simulation". Vol. 1 of *Proceedings of the IEEE Winter Simulation Conference*, Orlando, FL, USA, pp. 32-38.
- [9] Houshyar, A., and Nuila, V., 1993. "Required steps for successful design and implementation of simulation". *Computers and Industrial Engineering*, **25**, pp. 53-56.
- [10] Mahmoud-Jouini, S.B., Midler, C., and Garel, G., 2004. "Time-to-market vs. time-to delivery. Managing speed in Engineering, Procurement and Construction projects". *International Journal of Project Management*, **22**, pp. 359-367.

- [11] Juvinall, R. C., and Marshek, K.M., 1994. *Fondamenti della progettazione dei componenti delle macchine*. Edizioni ETS, Pisa, Italy.
- [12] Funaioli, E., Maggiore, A., and Meneghetti, U., 1998. *Lezioni di meccanica applicata alle macchine*, 2st ed., Vol. 1 and 2. Editrice Patron Bologna, Italy.
- [13] McConnel, P.V., 2007. "Software delivers faster time-to-part and reduced testing". *Reinforced plastic*, September, pp. 24-25.
- [14] Feng, Z.-Q., Joli, P., and Seguy, N., 2004. "FER/Mech – a software with interactive graphics for dynamic analysis of multibody system". *Advances in Engineering Software*, **35**, pp. 1-8.
- [15] Rio, G., Laurent, H., and Ble's, G., 2008. "Asynchronous interface between a finite element commercial software ABAQUS and an academic research code HEREZH++". *Advances in Engineering Software*, **39**, pp. 1010-1022.
- [16] Witt, W., 2006. *Simulation*, Vol. 13. of *Handbooks in Operations Research and Management Science*. North Holland publishing company, Amsterdam, Chap. 13, pp. 381-413.
- [17] Zach, M., and Siedl, D., 2007. "A New Method for Simulation of Machining Performance by Integrating Finite Element and Multi-body Simulation for Machine Tools". *CIRP Annals – Manufacturing Technology*, **56**, January, pp. 383-386.
- [18] Fleischer, J., Broos, A., Schoop, M., Wieser, J., and Hennrich, H., 2008 in press. "Life-cycle oriented component selection for machine tools based on multibody simulation and component life prediction". *CIRP Journal of Manufacturing Science and Technology*.

MICROSTRUCTURE PREDICTION IN AA6060 HOT PROFILE EXTRUSION

Lorenzo Donati

*Department of Mechanical, Aerospace, and Metallurgical Engineering,
University of Bologna, Italy
E-mail: l.donati@unibo.it*

Luca Tomesani

*Department of Mechanical, Aerospace, and Metallurgical Engineering,
University of Bologna, Italy
E-mail: luca.tomesani@unibo.it*

Marco Schikorra

*Institute of Forming Technology and Lightweight Construction,
Technische Universität Dortmund, Dortmund, Germany
E-mail: marco.Schikorra@iul.tu-dortmund.de*

A. Erman Tekkaya

*Institute of Forming Technology and Lightweight Construction,
Technische Universität Dortmund, Dortmund, Germany
E-mail: erman.Tekkaya@iul.tu-dortmund.de*

Abstract. *In the first part of the work an experimental procedure to investigate the evolution of recrystallization in aluminum alloys is presented and discussed. Several cups, obtained by means of inverse extrusion, were produced at different temperatures and process speeds. The specimens were analyzed in order to examine the grain size distribution. The coefficients for dynamic recrystallization models were obtained by regression analysis after thermo-mechanical FEM simulations of the experiments realized with the code Deform 3D.*

The specimens were then heated in a furnace and cooled in order to reproduce static recrystallization of the material. The grain distribution was examined and the coefficients for the equation for static recrystallization prediction were regressed, too.

In the second part of the work the extrusion of a round-shaped profile is described and the grain size distribution on the profile and on the billet rest are analyzed. The developed equations were applied to the rod extrusion and the simulated grain size distributions were compared with the experimental ones.

Keywords: *extrusion, aluminium alloys, FEM, microstructure evolution*

1. INTRODUCTION

In recent years the interest in microstructure prediction of metal formed products has seen an increasing attention, especially in the field of light alloys processing where the final mechanical properties strongly depend on the microstructure [1,2]. It is well known that grain size and particle precipitation strongly affect the product characteristics by different point of views: aesthetical appearance, product strength and deformability, fatigue resistance or corrosion. Usually, fine grain is desired, especially where high fatigue loads are applied or when the final product is exposed to a corrosion environment [3,4]. When extruding aluminum alloys the mechanical proprieties of the profile are of great interest, in particular strength and deformability of profiles for transportation industry. The mechanical properties of the profiles are mostly related to the microstructure evolution of the alloy during the whole production cycle, from billet casting to profile aging [5]. Coarse intermetallics precipitation, precipitate free zones (PFZ), grain size distribution and coarsening [6] are some examples of problems related to the aluminum alloy's metallurgical evolution that could lead to poor product proprieties at the end of the production sequence. Today, the desired microstructure is achieved mainly on the basis of process experience, but the use of process simulation is giving growing attention also in the everyday industrial production. Aluminum recrystallization has been one of the most investigated topic over the past years due to its specific behavior during hot deformation [7,8] but no paper provides reliable equations and coefficients for an effective FEM analysis. During the whole thermo-mechanical process, different metallurgical phenomena take place: static recrystallization SRX, dynamic recrystallization DRX (partially), geometric recrystallization GDRX, grain growth, and particle precipitation [9,10]. In the aluminum alloy processing, and in particular in the 6XXX series, the grains are strongly deformed to one principal direction due to the process strains, thus producing a high contraction of the grains in the other direction; when this thinning reaches the dimension of subgrains, each subgrain becomes a new grain. This phenomenon is usually called Geometric Dynamic Recrystallization (GDRX) and its mechanics are quite different from standard dynamic recrystallization where the nucleation of new grains occurs. Moreover, static recrystallization and grain growth act together when deformation ends (after the die exit) and the contribution of each distinct phenomenon can not be easily investigated.

With the aim of predicting grain size distribution on AA6060 extruded profiles, the study was divided into three different steps.

In the first step, a specially designed test has been developed in order to analyze the effect of recrystallization on the section of small cups obtained by a small scale inverse extrusion experiment. The cups were realized at different temperatures and ram speeds in order to embrace typical process conditions of the industrial extrusion and then immediately water-quenched. In the second step, the cups were heat-treated and the grain evolution was analyzed and measured again. The grain size distribution at each step was then used to regress equations for recrystallization prediction and such equations were implemented in the numerical analysis. In the third step the model was used in the simulation of an AA6060 rod profile and the results of the simulation were compared and discussed with the experimental grain size distribution measured on the extruded rests.

2. THE FEM ENVIRONMENT

The recrystallization prediction has been implemented inside the DeformTM 3D environment due to the availability of general equations for metallurgical evolution [11].

Deform is a thermo-mechanical coupled FEM code based on a Lagrangian approach: this method is useful for metallurgical prediction because it allows monitoring the evolution of strain, strain rate, and temperature during the different stages of the process.

In Deform, the recrystallization behavior of metals is classified into three broad categories [12]:

- Dynamic recrystallization (DRX): Recrystallization that occurs during deformation when the strain exceeds a critical strain.
- Static recrystallization (SRX): Recrystallization that occurs after deformation when the imposed strain is lower than a critical strain. Recrystallization begins in a nuclei-free environment.
- Grain growth, which occurs before recrystallization begins or after recrystallization is completed. The driving force is the reduction of grain boundary energy.

In this work, according to the above classification and to the mechanisms taking place in aluminum alloy deformation, the adopted DRX model will take into account all recrystallization mechanisms that are present during deformation throughout the experimental procedure. Thus, geometric dynamic recrystallization, dynamic recrystallization, and continuous dynamic recrystallization will be modeled as a unique effect composed by different contributions. Moreover, as it will be explained below, the analytical models for the recrystallization prediction (and in particular static recrystallization) available inside the FEM environment are actually upper limited to the initial grain dimension; in order to overcome such limitations, only the grain growth model (coupled with the dynamic recrystallization model) will be considered in this study with the aim to include also the static recrystallization effect.

Dynamic Recrystallization (DRX) This mechanism occurs during deformation when the strain exceeds a critical strain, the driving force being the removal of dislocations. Experimental data must be collected under various strain, strain rate, and temperature conditions. The critical strain ε_c is usually a fraction of the strain ε_p at which the flow stress reaches its maximum (and flow softening due to dynamic recrystallization starts).

$$\varepsilon_c = a_2 \varepsilon_p \quad (1)$$

The value of ε_p is determined experimentally and is usually a function of strain rate, temperature, and initial grain size:

$$\varepsilon_p = a_1 d_0^{n_1} \dot{\varepsilon}^{m_1} \exp(Q_1 / RT) + c_1 \quad (2)$$

An Avrami equation is also used to describe the relation between DRX volume fraction and the effective strain:

$$X_{drx} = 1 - \exp \left[- \beta_d \left(\frac{\varepsilon - a_{10} \varepsilon_p}{\varepsilon_{0.5}} \right)^{k_d} \right] \quad (3)$$

in which $\varepsilon_{0.5}$ denotes the strain for 50% recrystallization:

$$\varepsilon_{0.5} = a_5 d_0^{n_5} \dot{\varepsilon}^{m_5} \exp(Q_5 / RT) + c_5 \quad (4)$$

The DRX grain size is expressed as:

$$d_{drx} = a_8 d_0^{h_8} \varepsilon^{n_8} \dot{\varepsilon}^{m_8} \exp(Q_8 / RT) + c_8 \quad (5)$$

where if $d_{rex} > d_0$ then $d_{rex} = d_0$.

Grain Growth Grain-growth modeling is performed for the material in the strain-free state; i.e., before hot working or after recrystallization is completed. For this purpose, the classical phenomenological grain growth relation is employed. The kinetics is described by equation:

$$d_{ggave} = \left[d_{ave}^{m_9} + a_9 t \exp\left(-\frac{Q_9}{RT}\right) \right]^{1/m_9} \quad (6)$$

Where d_{ggave} denotes the average grain size after growth, a_9 and m are material constants, and Q_9 is the activation energy.

Temperature Limit The temperature limit is the lower bound of all grain evolution mechanisms. Below this temperature, no grain evolution is computed.

Average Grain Size The mixture law is employed to calculate the recrystallized grain size for incomplete recrystallization:

$$d_{ave} = X_{rex} d_{rex} + (1 - X_{rex}) d_0 \quad (7)$$

The nomenclature of the above formulation is: T – Temperature (°K); R - Gas constant (8.314472 J/(K · mol)); t – Time (sec); d_0 - Initial grain size (μm); d - Current grain size (μm); d_{rex} - Recrystallized grain size (μm); d_{ave} - Average grain size (μm); X –Volume fraction of recrystallization (X_{drx} dynamic, if $X=0$ no recrystallization occurred, if $X=1$, 100% recrystallization occurred); ε - Effective strain; ε_c - Critical strain (strain above which dynamic recrystallization occurs); ε_p - Peak strain (strain at peak value of effective stress); $\varepsilon_{0.5}$ - Strain for 50% recrystallization; $\dot{\varepsilon}$ - Effective strain rate (sec^{-1}); $t_{0.5}$ - Time for 50% recrystallization (sec); a_{1-10} - Material data (experimental coefficients); h_{1-8} n_{1-8} m_{1-9} - Material data (grain-strain-strain rate experimental exponent); Q_{1-8} - Material data (Activation Energies); β_d - Material data experimental coefficient; k_d - Material data experimental exponent.

Table 1. Dynamic recrystallization regressed coefficient.

<i>Alloy</i>	<i>Temperature</i>	n_g	m_g	a_g
AA6060	250° C	-0.364	-0.213	1.93E+15
	350° C	-0.985	-0.105	7.22E+12
	450° C	-0.722	-0.084	1.34E+11
	550° C	-0.420	0.046	8.26E+09

3. RESULTS AND DISCUSSION

First Step-Laboratory Test

In the first step a laboratory scale testing procedure was defined in order to evaluate the effect of different processing conditions on the evolution of the grain size distribution after deformation of an AA6060 alloy provided in standard industrial homogenization conditions (chemical composition in wt% Si 0.43; Fe 0.21; Cu 0.02; Mn 0.03; Mg 0.48; Zn 0.02; Cr 0.0012; Ti 0.01; Al remaining). An inverse extrusion equipment was designed (figure 1), thus producing eight cups with decreasing wall thickness at different temperatures (250°, 350°, 450° and 550° C) and ram speeds (0.1 and 5 mm/sec).

The testing parameter range was chosen in order to obtain local conditions of strain (0-3.88), strain rate (0-5 s⁻¹), and the temperature inside the specimens was typical to industrial extrusion processes. The specimen and the testing tools were placed inside a resistance oven in order to reach the temperature constant during the press stroke; the specimen temperature was controlled by means of a contact thermocouple. Immediately after the deformation the specimens were water-quenched in order to fix the metallurgical structure. The deformed specimens and an undeformed one were then embedded and analyzed by means of polarized optical microscopy after an electrolytic etching in a Barker's reagent; finally, the grain size distributions were measured and recorded. Besides, the whole testing plan was simulated by means of the Deform code in order to obtain the distribution of strain and strain rate inside each specimen. By a comparison of the measured grain size with local forming conditions the experimental coefficients of the dynamic recrystallization model were regressed as follows: $a_2=0.05$, $\varepsilon_{0.5}=0.15$, $a_5=0.15$, $\beta_d=1$ $k_d=1$, $a_{10}=1$, n_g , m_g , a_g (table 1), $h_5=n_5=m_5=Q_5=0$.

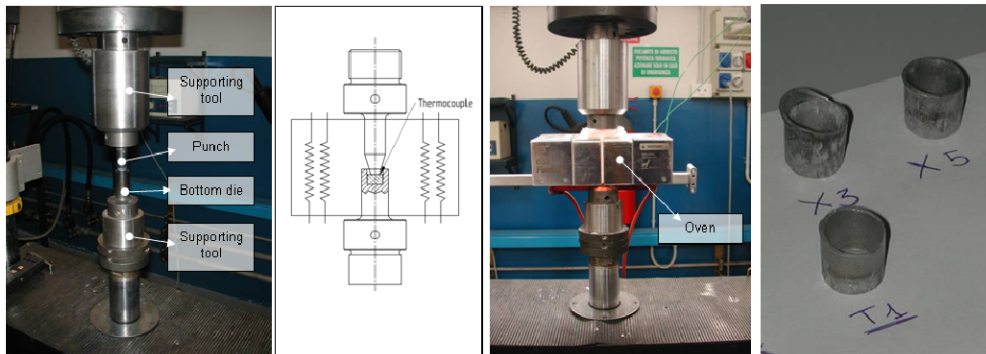


Figure 1. Inverse extrusion equipment.

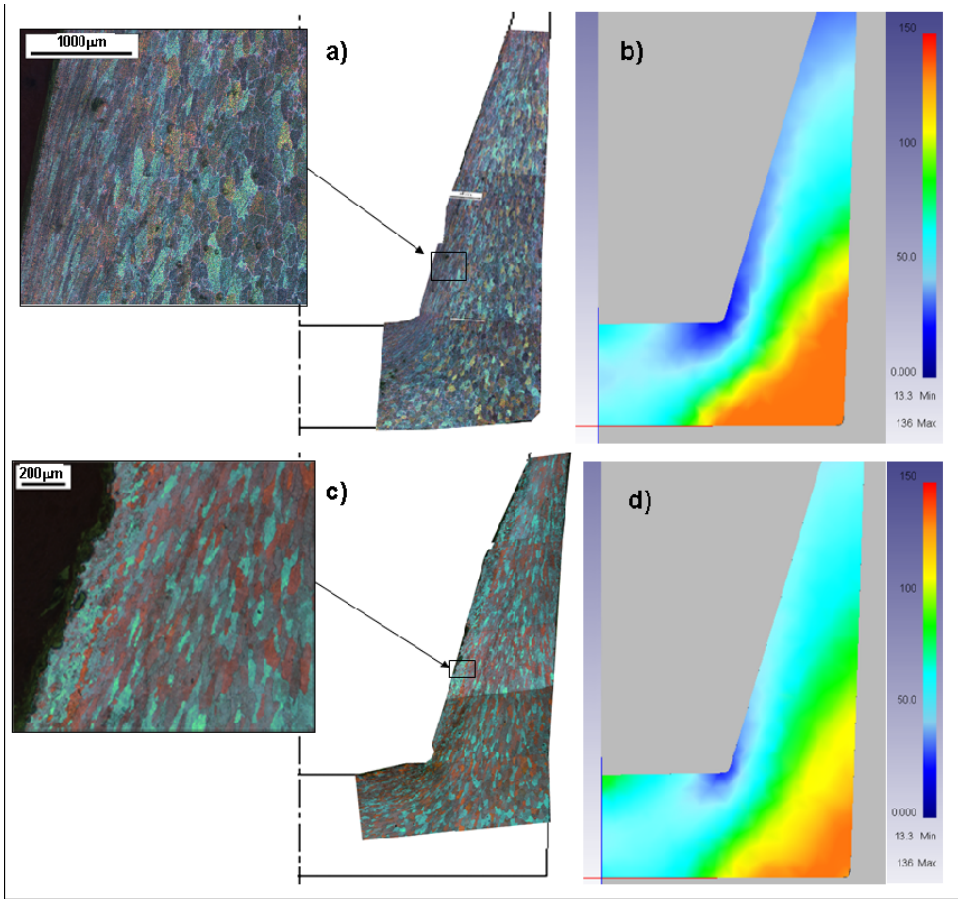


Figure 2. Temperature=350° C, speed=0.1 mm/sec: a) Optical Micrograph; b) Simulated grain size; Temperature=550° C, speed=0.1 mm/sec: c) Optical Micrograph; d) Simulated grain size.

A more detailed description of the experimental procedure, FEM simulations, and coefficient regression is reported in [13].

In the base material the microstructure consisted of well-developed and equiaxed coarse grains with an average size of approx. 137 μm with large constituent particles along the grain boundaries, while other intermetallic compounds are observed inside the grains. Instead, in figure 2a the typical grain size distribution inside the deformed specimens is shown. Widely elongated grain formations (only deformed) were found in proximity of the punch where a maximum strain of 2.5-3.88 was estimated, while more equiaxed grains (hardly deformed at all) were observed near to the die surface where a strain below 0.8 was found. A general decrease of grain thickness was noticed moving from the die to the punch: a thickness of 20-50 was observed in proximity of the punch depending on the testing conditions, while of 80-110 μm next to the die. This decrease can be explained by the deformation of the material in axial direction, while only in few locations the geometric

dynamic recrystallization effect (GDRX) produced new grains without any nucleation of new grains. In fact, the grains are strongly deformed along the extrusion direction and when the thickness of the grain reaches the subgrain dimension a new grain is formed.

By comparing the experimental measurements with the FEM-predicted grain sizes distributions (fig. 2b) throughout the whole specimens section, a mean error of 12% with a maximum deviation of 57% was found, thus giving an interesting agreement with the experimental results, as reported in detail in [13]. The greatest differences are usually found at the lowest temperature and in the upper section of the cups where a tools misalignment can provide inhomogeneous deformation and, consequently, a non-uniform wall thickness distribution in the specimens.

Second Step-Heat Treatment

In the second step the eight specimens were heat-treated in a laboratory furnace at 550° C for 30 min, then air-cooled to room temperature, and finally heated again to 180° C for 10 hrs. Following the same procedure of the first step, the specimens were sectioned, embedded, and the grain dimensions were analyzed.

Equiaxed grains were now found inside the entire specimens (fig. 3), thus indicating a complete recrystallization of the material and a general increase of grain dimension: in particular, the grain sizes were always bigger in the specimens where a lower testing speed was used. Moreover, with the increase of the testing temperature an increase of grain dimension was found, and in particular when the testing temperature exceeded 450°C an abnormal peripheral grain growth (PCG) was observed (fig. 3-c).

When PCG was absent (i.e. fig.3-a), a grain dimension of 140-170 μm was found next to the punch, while next to the die bigger grains were usually found (180-250 μm), again depending on the testing conditions during deformation; on the other hand, when PCG occurred in the area of high strain, an abnormal grain dimension (up to 1500 μm) was found; at lower temperatures the abnormal growth is limited to high strain areas (up to 2), while with increasing temperature the PCG area is extended to lower strains (up to 0,6). It is interesting to notice that, also when PCG is absent, grain dimensions are bigger than the initial one (137 μm).

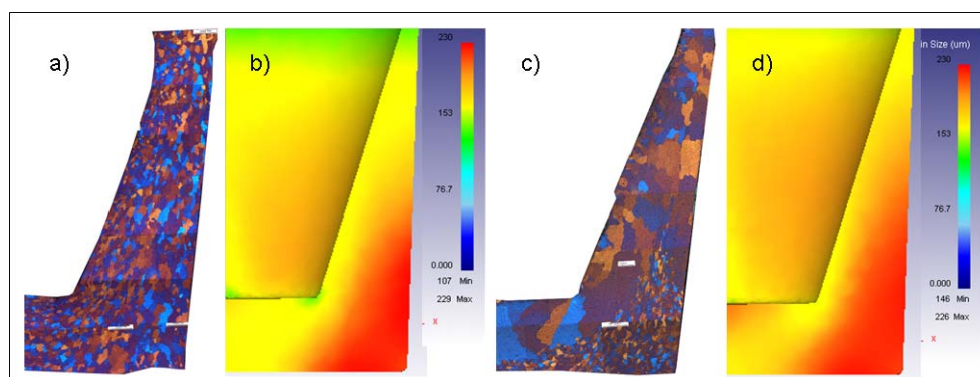


Figure 3. Specimens after heat treatment: a) 350°C-0.1 mm/sec, b) 350°C-0.1 mm/sec simulated, c) 550°C-0.1 mm/sec, d) 550°C-0.1 mm/sec simulated.

Besides, the same thermal cycle was simulated with FEM and the parameters of the equation (6) were regressed by comparison with experimental results ($a_0=14$, $Q_0=47600$ and $m_0=0,8$). The simulated grain size distribution obtained with the grain growth model is reported in figure 3 (b and d). The developed model describes the increasing in grain size with testing speed and temperature very well: a mean global error of 14% with a peak of 49% in the prediction was found, but in the calculation of this error the locations with PCG were excluded. Some limitations arise with the utilized grain growth model (6): first of all, the applied model does not consider the amount of energy stored in the material in the form of dislocations, so all grains growth independently from their deformation history; secondly, PCG could not be computed by such model because coarsening is a phenomenon mainly related to the maximum strain reached in the material and in an analytical model formulation no direct effect of the maximum strain is considered.

Third Step-Extrusion of a Rod Profile

The experimental investigations were carried out using a simple axisymmetric die shape in order to guarantee a symmetric microstructure distribution. The profile, a round rod of 36 mm diameter (extrusion ratio 14.7) made of AA 6060 alloy, was produced at 0.3 mm/sec ram speed under strict observation of the process conditions. For the development of a microstructure (recrystallization and grain grow) the temperature is one of the most important influencing parameters. In the experiment the temperatures were strictly monitored by thermocouples during the whole extrusion process, starting from initial heating to process handling, extrusion, and cooling down. The material temperature history is schematically reported in figure 4, while the monitored data during the press stroke are shown in figure 5. As reported in figure 4, the billet reached a temperature of 550° C inside the oven then decreased to an exiting temperature of the profile of approximately 400° C. This temperature history, which does not evidence any rise by deformation, is related to both the quite low extrusion ratio and the low temperatures of container and die (450° C and 380° C respectively) that cooled down the material temperature flowing through the die. At the same time, the die temperature increased. At the end of the extrusion stroke it has to be distinguished between temperature of the profile and temperature of the billet. Due to the time needed to extract the billet from the container the billet did not cool down immediately, while the temperature of the exiting profile started to decrease directly after the die exit. Both were cooled with ambient air to room temperature.

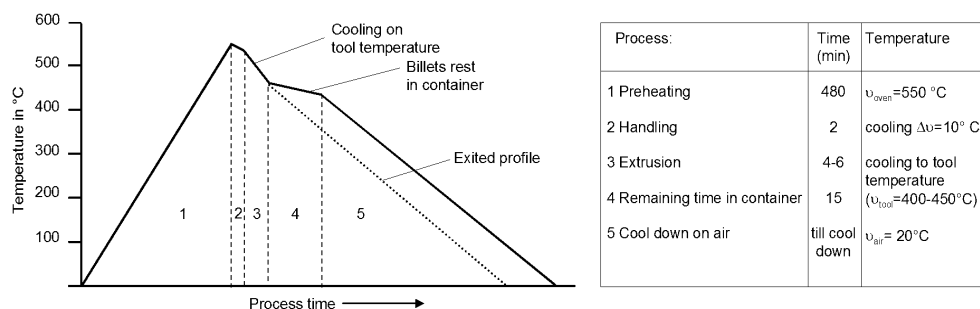


Figure 4. Billet and profile temperature over process time, schematic

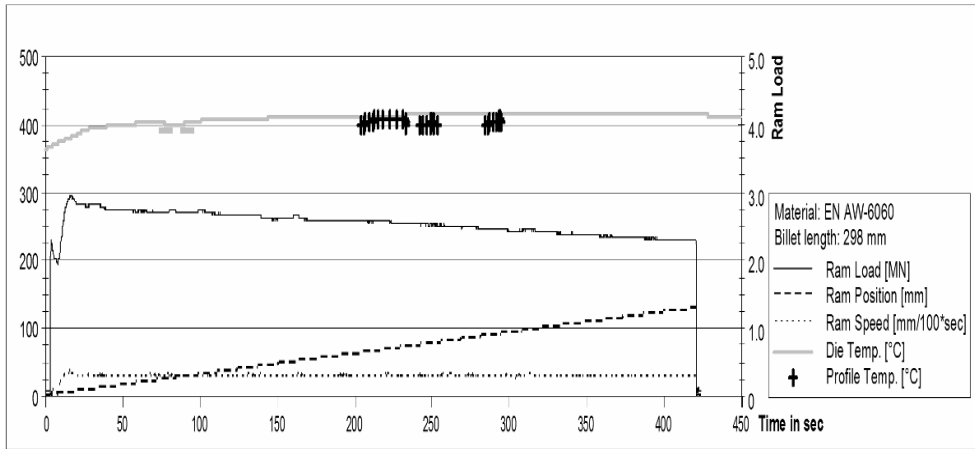


Figure 5. Monitored process conditions during the extrusion of AA6060 billets

Because an immediate quench of the billet rests after extrusion was not possible the analysis contains recrystallization during the forming step as well as static recrystallization and grain growth when the material remains at high temperature in the container and during the cooling down phase.

As it can be seen in the curves of figure 5, after the initial upsetting of the billet from $\phi 138$ to $\phi 146$ container diameter, the punch load reached a maximum value of 3 MN. For an analysis of the material flow and the grain structure not only in the profile, but also in the press rest, the extrusion was stopped after a punch travel of approximately 40-50 percent of the billet's length (Ram stroke = 130 mm) in order to have a larger area for the analysis of the material grain size at different strain conditions.

Then, the billet was prepared for metallurgic investigations; it was sawed next to the symmetry plane, machined and polished, then etched at room temperature for 5-8 minutes with 90 ml H₂O, 15 ml HCl, and 10 ml HF based on Flick procedure.

Figure 6 shows the macro and micro sections of the billets after the etching procedure in the symmetry plane. As shown in [14], it can be observed on the macro etching that the microstructure differs significantly regarding the grains' position in the billet rest. Aiming at a complete analysis of the material deformation path during extrusion, figure 6 reports both the macro images of the complete billet and micro images along several deformation paths. Four different zones develop inside the rest which can be distinguished by their etching response like grain pattern, grayscale level, and reflection. Material flow lines are reported in order to separate the major zones of the deformation process.

Four different zones are marked: zone I is the dead metal zone (DMZ), in this region the material has almost no movement during the ram stroke. It is exposed to high temperatures and high hydrostatic pressure, but small strain and negligible strain rate occurred (except for the upsetting phase). The microstructure in this area (A1) is dominated by a small reduction in size of the initial microstructure; perfect equiaxed grains are evidenced and an average grain size equal to 110 μm is found.

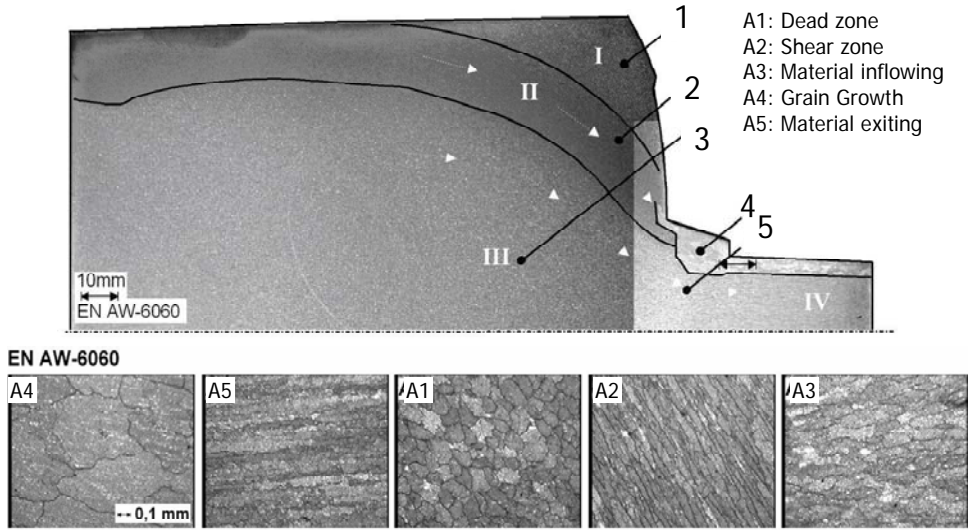


Figure 6. Macro and micro analysis of the extruded rest (AA6060).

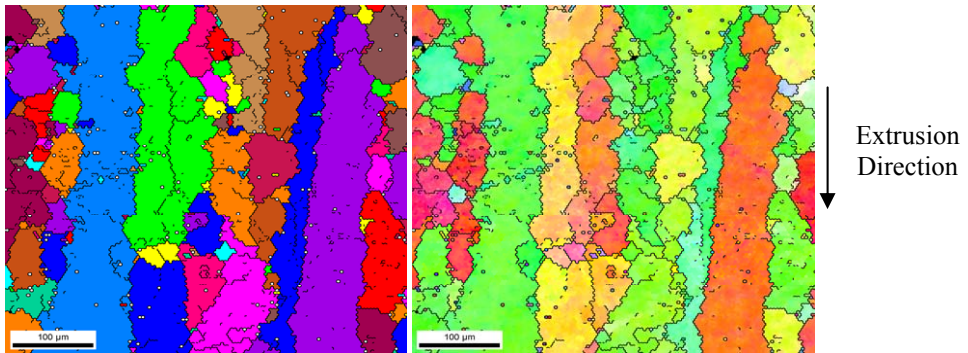


Figure 7. EBSD analysis of A5 location.

Due to the low level of plastic strain no recrystallization occurred, only small grain elongation took place. Furthermore, no grain growth will appear according to the lack of recrystallization. Zone II is the shear intensive zone (SIZ) of deformation and it outlines the DMZ. The grains in this area undergo high shear strain deformation caused by the boundary conditions: still material is located in the external part due to the friction effect (DMZ), while high speeds (sometimes higher than ram speed) move the material in the inner part of the billet. Strain and strain rates are high, but not the highest yet (see figure 8), as better described later. The grains (A2) are orientated as the material flows and elongates to long narrow banded stripes; in relation to this high elongation the grains of the shear intensive zone have an average length of 320 μm and a width of 54 μm . If an equivalent grain (in term of section area) of round shape is computed, a mean diameter of 120 μm is found.

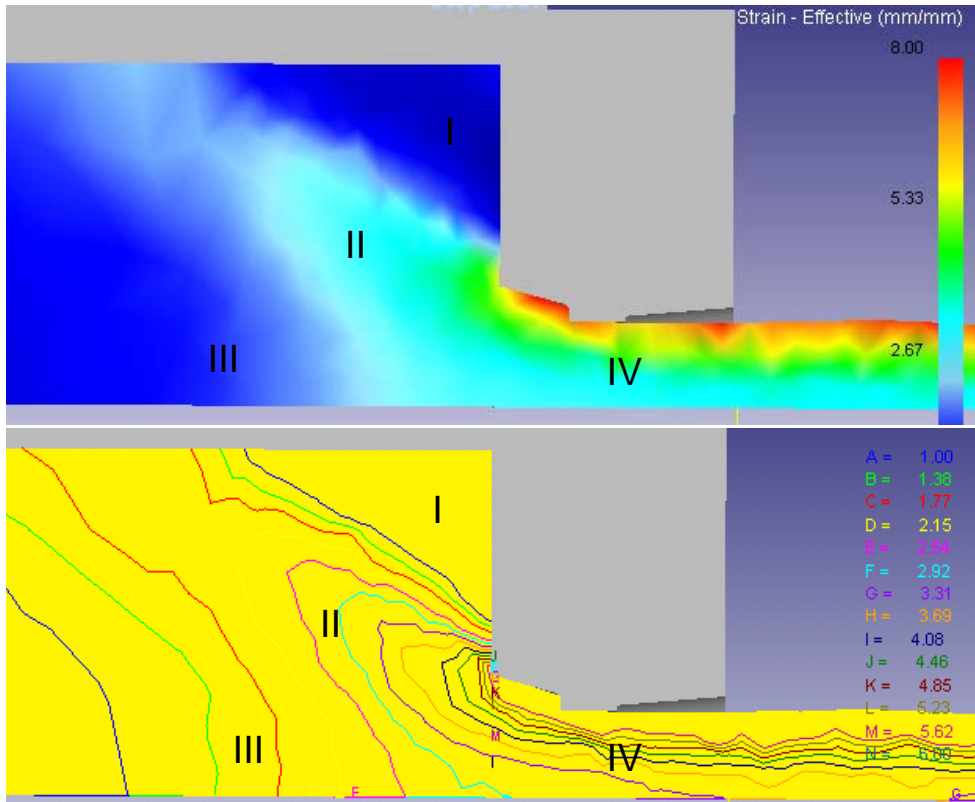


Figure 8. Distribution of strain in the billet rest.

Due to the high degree of deformation and even high strain rates this area shows relevant geometric recrystallization, but unexpectedly negligible grain growth is found, probably due to low temperatures. In contrast to SIZ zone, the material in the main deformation zone III (MDZ) flows directly into the die orifice. The grains do not deform or deform only slightly within the billet. They undergo a translation movement that leads to a plastic deformation only near the die opening. The grain structure (A3) is mainly influenced by the initial grain size after homogenization (average grain thickness=100 μm). The material undergoes high temperatures and high hydrostatic pressure and strain and strain rate increase with decreasing distance to the orifice. A slight geometric dynamic recrystallization effect in this zone usually appears at short distance from the die exit (A5) as evidenced by EBSD analysis of figure 7: recrystallized grains of 50 μm diameter are found inside bigger elongated grains of 80-100 μm thickness.

The fourth zone is the exiting profile zone (EPZ), which is the result of all the above mentioned influences. In this zone the microstructure can differ drastically, depending on the recrystallization history in relation to the position in the profile. A complicated mixture of high strains, strain rates, and temperature, but also cooling rates occur near to the surface. Speeds are quite high, so the temperatures rise to their maximum values; the same happens to the strain that reaches the highest values in the bearing.

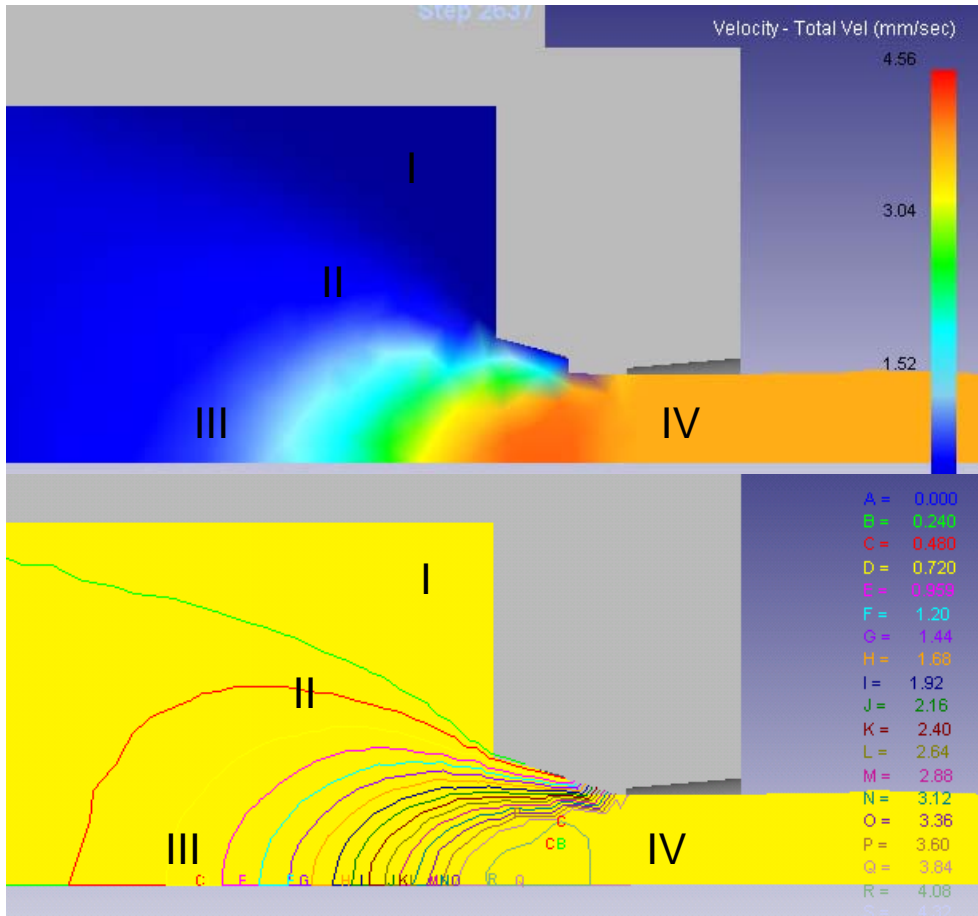


Figure 9. Distribution of material velocity in the billet rest.

Analyzing the coarse grains just below the surface of the die orifice (A4), these show a grain size of around 1000-1300 μm diameter. A profound material recrystallization process and, furthermore, an abnormal grain growth took place because there are no more visible deformed grains here. This recrystallization process was first geometric-dynamic during the deforming stroke, then static growth occurred after die exit. The same effect was analyzed on the cups in the second step of this work where a mixed effect of strain and temperatures was evidenced as being responsible for the PCG formation.

The parameters optimized in the first two steps were then applied to the extrusion of the rod. The monitored process conditions were used as input settings for the FEM simulation. A stroke of 130mm at a ram speed of 0.3 mm/sec was computed while the friction was considered as sticking, as determined in a previous work by the authors [15]. After the deformation stroke, the billet somewhat decreased its temperature by remaining for 15 minutes in contact with the container and then air cooling was performed and computed. In figure 8 and 9, the distribution of strain and material velocity in proximity of the final deformation stroke are plotted: again, the four different zones can be clearly

detected, depending on the amount of strain and speed. In the billet, zone I (Dead Metal Zone, DMZ) is characterized by strains lower than 1, the material is highly hydrostatically compressed, but the material velocity is so low that it can be considered almost still; in zone II (Intensive Shear Zone, SIZ) a high strain is evidenced (1,7-4,5), but low flow rates (0.2-0.4 mm/sec) are present; in zone III (main deformation zone, MDZ) the material has low strains (below 1,2) and the material speed is similar to the ram speed (0.3 mm/sec).

Finally, in zone IV (Exiting Profile Zone, EPZ) a homogeneous high speed (4.1 mm/sec) is evidenced, while two regions of strain are present: a superficial one with high strains (around 8) and an inner one with lower values (around 2.7). As to the temperature, the billet was preheated to 500° C and due to the lower die temperature (380° C) the exiting profile temperature results to be quite low (380-400° C).

In figure 10 the simulated average dimension of the grain is plotted: in zone I (point 1) a grain size around 105 μm is computed, in close agreement with the 110 μm value of the experimental measurement; in points 2 and 3 (zone II and III) the same grain size prediction is obtained (116μm).

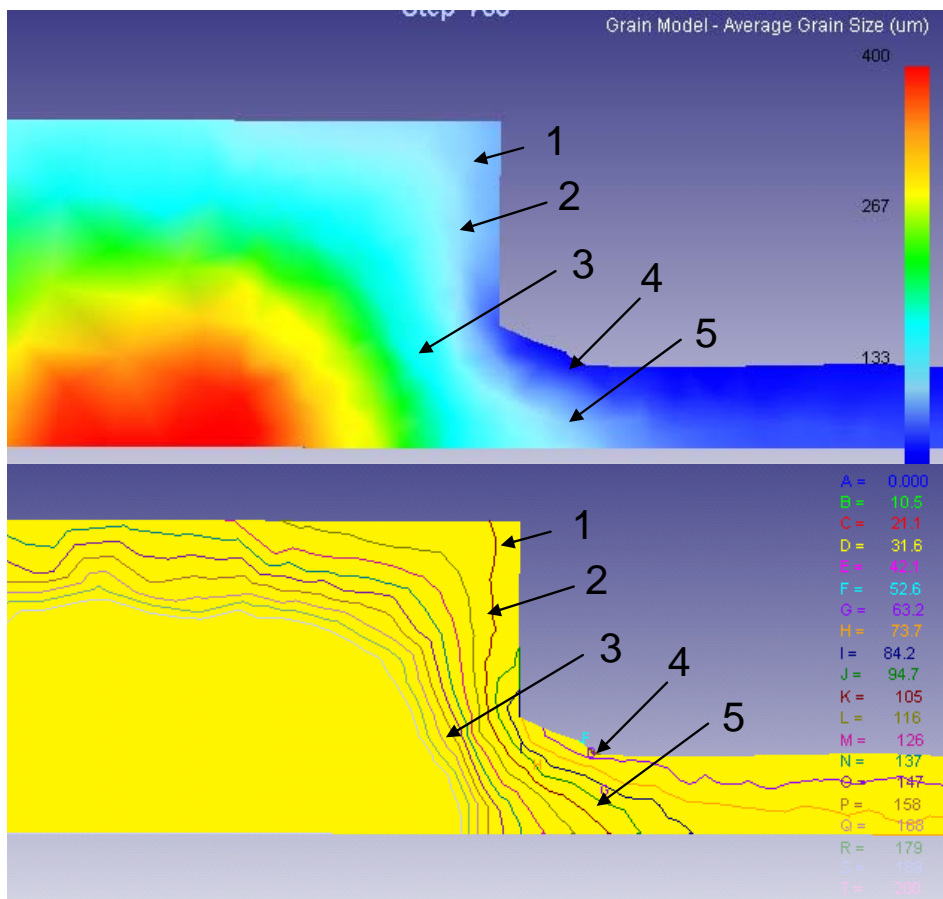


Figure 10. Distribution of average grain size in the billet rest.

In 2, the grains are highly distorted because static recrystallization did not occur and, in this condition, it is difficult to estimate the grain size: if minimum grain thickness is computed, a big difference between computed value (116 μm) and experimental (54 μm) is found; otherwise, if we consider the equivalent grain area, the experimental dimension is around 120 μm , thus providing a good agreement with measurements. Besides, in the area on the left of point 3 an unacceptable prediction of grain dimension is made: a grain of 190 μm is evaluated instead of the 120 μm (equivalent) measured. This discrepancy can be explained by the adopted definition of a grain growth model that does not discerns points where high deformation occurred from points where small strains are present. In fact, at point 3 the code considers grain growth, whereas only few dislocations are present in the material so as to limit the growth of the grain to a minimal extent.

Finally, in zone IV (points 4 and 5) a prediction of 120 and 70 μm is made with respect to the 1000 and 50-100 μm of the experimental measurement. According to the observation that at temperatures above 450° C PCG is formed if the strain exceeds a critical value and that with decreasing temperatures higher strains are obtained, it is possible to notice that a peripheral coarse grain was formed in the extruded rest, where strains reached a value higher than 5. In this direction, the grain growth model has to be modified and adapted to the dependencies of the phenomena to strain and deformation energy.

4. CONCLUSIONS

In this work an experimental methodology to evaluate the prediction of recrystallized structures in aluminum extrusion was presented and validated.

- 1) Inverse extrusion tests were realized in order to obtain the widest set of thermal and mechanical history conditions for microstructure development. In particular, the effects of geometric dynamic recrystallization (GDR) and static recrystallization (SRX)/grain growth were separately analyzed and the peripheral coarse grain (PCG) phenomenon was investigated.
- 2) The whole experimental plan was simulated by FEM analysis with Deform 3D in order to derive the local conditions of strain, strain rate, and temperature to be correlated with the evaluated microstructures.
- 3) A complete set of coefficients was regressed for the available microstructure evolution models inside the code environment.
- 4) The obtained models were applied to the real extrusion of a round profile and a comparison between experimental measurements and simulation results was performed. The simulated results were in very good agreement with experimental data, except in zones where peripheral coarse grain and grain growth appeared. Here, a further investigation effort and specific modeling equations are required.

Acknowledgements

This work was carried out with the financial support of the MIUR (Italian Ministry for Research and Innovations) and the Transregional Collaborative Research Center/TR30 funded by the German Research Foundation (DFG). The authors would like to thank Dr. Jie Zhou and Dr. Joanna Duszczek for the cooperation during the extruded cup analysis.

REFERENCES

- [1] T. Sheppard, 2006. "Prediction of structure during shaped extrusion and subsequent static recrystallisation during the solution soaking operation", *Journal of Materials Processing Technology* vol.177, pp. 26–35.
- [2] A.R. Bandar, S. R. Claves, J. Lu, K. Matous, W. Z. Misiolek, A. M. Maniatty, 2004. "Microstructural Evaluation of 6xxx Aluminum Alloys for Computer-Simulated Texture Prediction", *Aluminum Extrusion Technology Seminar ET 2004, Orlando*, vol 1, pp169-176.
- [3] B. Dixon, 2000. "Extrusion of 2xxx and 7xxx alloys", *Aluminum Extrusion Technology Seminar, Chicago*, vol1, pp281-2940.
- [4] T. Sheppard, 1996 "Development of structure recrystallization kinetics and prediction of recrystallised layer thickness in some Al-alloys", *Aluminum Extrusion Technology Seminar, Chicago*, vol1, pp163-170.
- [5] J.M.C. Mol, J. van de Langkruis, J.H.W. de Wit and S. van der Zwaag, 2005. "An integrated study on the effect of pre- and post-extrusion heat treatments and surface treatment on the filiform corrosion properties of an aluminium extrusion alloy" *Corrosion Science*, Volume 47, Issue 11, pp 2711-2730
- [6] F. J. Humphreys, M. Hatherly, *Recrystallization and Related Annealing Phenomena*, Pergamon Press Inc, Oxford, 1995 ISBN 978-0080418841
- [7] S. Gourdet, F. Montheillet, 2000. "Experimental study of the recrystallization mechanism during hot deformation of aluminium" *Materials Science and Engineering A: Structural Materials: Properties, Microstructure and Processing*, v 283, n 1-2, May, pp 274-288
- [8] J. G. Byrne, *Recovery, Recrystallization, and Grain Growth*, (New York: MacMillan, 1965), 93-109.
- [9] R. D. Doherty, D. A. Hughes, F. J. Humphreys, J. J. Jonas, D. Juul Jensen, M. E. Kassner, W. E. King, T. R. McNelley, H. J. McQueen and A. D. Rollett, 1997. "Current issues in recrystallization: a review" *Materials Science and Engineering A*, Volume 238, Issue 2, Pages 219-274
- [10] T. Pettersen, B. Holmedal, E. Nes, 2003 "Microstructure development during hot deformation of aluminum to large strains" *Metallurgical and Materials Transactions A: Physical Metallurgy and Materials Science*, v 34, n 12, December, pp 2737-2744
- [11] J. Fluhrer, 2006. *DEFORMTM 3D User's Manual Version 6.0 Scientific Forming Technologies Corporation*
- [12] G. Shen, S.L Semiatin, and R. Shivpuri, 1995. "Modeling Microstructure Development during the Forging of Waspaloy", *Metallurgical and Materials Transactions A*, 26A, 1795-1803.
- [13] L. Donati, J. Dzwonczyk, J. Zhou, L. Tomesani, 2008. "Microstructure prediction of hot-deformed aluminum alloys" *Key Engineering Materials* Vol. 367, pp. 107-116;
- [14] T. Sheppard, 1984. "Metallurgical Aspects of Direct and Indirect Extrusion", *Proc. of the 4th Aluminum Extrusion Technology Seminar*, 107-124
- [15] M. Schikorra, L. Donati, L. Tomesani, M. Kleiner, 2007. "The role of friction in the extrusion of AA6060 aluminum alloy, process analysis and monitoring", *Journal of Materials Processing Technology*, 191 pp. 288–292.

MISURA SPERIMENTALE E PREVISIONE DELLE FORZE DI TAGLIO, PER MATERIALI DALLE CARATTERISTICHE MECCANICHE VARIABILI

Giovanni Tani

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: giovanni.tani@mail.ing.unibo.it*

Claudio Mantega

*Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: Claudio.mantega2@unibo.it*

Abstract. *Questo lavoro tratta la previsione delle forze che si sviluppano nel processo di fresatura, con frese a codolo, su pannelli di materiale compostino a fibra legnosa: MDF (Medium density fiberboard). Un modello analitico per la previsione delle forze di taglio, sviluppato per il taglio dei metalli è stato utilizzato opportunamente tarato sul processo di taglio di pannelli in MDF. Il modello analitico è in grado di calcolare in tempi brevi le tre componenti della forza di taglio. Le forze di taglio sono modellate come funzioni delle condizioni di taglio, della geometria dell'utensile dei parametri di taglio per l' MDF, e sono espresso in funzione dell'angolo di rotazione dell'utensile, oppure del tempo. Il modello è stato implementato in Matlab-Simulink al fine di realizzare un tool di calcolo fruibile a livello aziendale.*

Ulteriori modelli analitici sono presenti nella letteratura delle lavorazioni del legno, alcuni di essi sono stati presi in considerazione, analizzati e alcune delle procedure di stima delle forze di taglio, sono state sviluppate e confrontate con il modello derivante da letteratura dei metalli.

Keywords: *pannelli MDF, previsione delle forze di taglio, fresatura.*

1. INTRODUZIONE

La previsione delle forze di taglio ha un ruolo importante, nella ottimizzazione dei parametri di taglio, così come nel controllo dell'usura e sulla qualità della superficie lavorata. Nella progettazione delle macchine utensili, la previsione delle forze di taglio è utilizzata per il dimensionamento dei componenti di macchina e delle strutture [1].

Diversi modelli matematici sono presentati in letteratura per valutare le forze di taglio; alcuni di essi prendono in considerazione parametri sperimentali che devono essere ottenuti da prove di taglio.

In questo articolo diverse tecniche e modelli sono riportati in modo da valutare le forze di taglio sviluppate durante la fresatura di pannelli in MDF.

I principali lavori meccanica del taglio del legno sono stati effettuati nel 1950 da Kivimaa (1950), Franz (1958) e McKenzie (1960).

Per studiare processo di fresatura sotto ipotesi di taglio ortogonale, Kivimaa [2] e Pahlitzsch / Jostmeier [3] suggeriscono modelli di taglio sulla base del Eq. (1):

$$F_c = k_c \cdot h_m \cdot b \quad (1)$$

$$k_c = w_0 + \frac{K}{h_m} \quad (2)$$

Dove , F_c è la forza di taglio media, h_m è lo spessore di truciolo, b medio larghezza di taglio, e k_c è un coefficiente di taglio.

Il coefficiente k_c è chiamato anche “ k_c -value” ed è largamente utilizzato in letteratura[4-6] e può essere assunto come una funzione iperbolica dello spessore medio di truciolo.

Tröger, and Dietz definirono che la relazione tra k_c value e h_m può essere scritta come in Eq. (2), dove I parametri K (che figurano in Eq.(2)) e w_0 sono determinati sperimentalmente.

Scholz raccolse diversi risultati delle analisi di dati sperimentali disponibili da pubblicazioni [4-6], ha anche verificato che il coefficiente di dipendenza lineare e l'inverso dello spessore di truciolo è ben definita per spessori maggiori di 0,2 millimetri .

A differenza di taglio dei metalli, in cui il processo è dominato dalla deformazione plastica di taglio, la lavorazione dei pannelli in MDF, è dominata da fratture e compressione delle fibre legnose .

Alcuni modelli di forze di taglio per taglio dei metalli (taglio sulla base di plastica) non sono utili per il taglio del legno. Tuttavia, grazie al comportamento sostanzialmente isotropo dei pannelli, alcuni modelli, possono essere trasferiti al caso in oggetto.

Nella letteratura sulla lavorazione dei metalli, è ben noto il modello di Altintas e Lee [7]; questo modello definisce il calcolo delle forze di taglio grazie a sei specifici parametri di taglio (K), valutati da tests di lavorazione.

In questo processo lavoro è stato affrontato utilizzando il di modello Altintas [7, 8], che è stato implementato in Simulink, al fine di realizzare un sistema computazionale di utenza aziendale.

Prove meccaniche sono stati quindi effettuati al fine di valutare i coefficienti di taglio (coefficienti K_{ij} presenti nelle Eq.(3)) richiesti dal modello.

Tali coefficienti di taglio sono determinati, per la particolare casistica di taglio, tramite prove di lavorazione, su macchina strumentata, viene di seguito esposta la procedura sperimentale, e la manipolazione dei risultati per pervenire alla definizione dei coefficienti.

Saranno quindi esposti i confronti tra le forze simulate e quelle misurate durante le prove.

È stata studiata e sviluppata anche la procedura di Kivimaa-Pahlitzsch, tale modello è in gradi di fornire l'entità della forza media di taglio. Infine viene riportato un confronto tra le due modellazioni sviluppate.

2. IL MODELLO DELLE FORZE DI TAGLIO

Nel modello analitico di Altintas [7, 8], la forza totale agente sul punto P_z sul tagliente dell'utensile, può essere scomposta nelle sue tre componenti rispetto ad un sistema di riferimento (tangenziale, assiale e radiale, Fig.1a) secondo le Eq.(3):

$$\begin{aligned} dF_t &= K_{te} dS + K_{tc} h(\phi, \kappa) db \\ dF_r &= K_{re} dS + K_{rc} h(\phi, \kappa) db \\ dF_a &= K_{ae} dS + K_{ac} h(\phi, \kappa) db \end{aligned} \quad (3)$$

dove dS è la lunghezza infinitesimale di un segmento del tagliente, db è la proiezione di un segmento di tagliente lungo la direzione della velocità di taglio e h è lo spessore di truciolo. Il punto P_z sul tagliente ha quota assiale z , distanza radiale $R(z)$ sul piano XY, e angolo di immersione $\kappa(z)$ e angolo di ritardo $\psi(z, \kappa)$, Fig.(1a).

L'angolo di immersione assiale è definito come l'angolo tra l'asse dell'utensile e la normale al tagliente nel punto P_z .

Nel modello generalizzato la relazione tra lo spessore di truciolo e l'avanzamento per dente s_t è riportata nella Eq.(5). l'angolo di immersione radiale ϕ_j varia in funzione dell'angolo di rotazione, ϕ , secondo l'Eq.(6).

$$db = dz \cdot \text{sen}(\kappa) \quad (4)$$

$$h = s_t \cdot \text{sen}(\phi_j) \cdot \text{sen}(\kappa) \quad (5)$$

$$\phi_j(z) = \phi + \sum_{j=1}^N \phi_{pj} - \psi(z) \quad (6)$$

Grazie ad un modello parametrico e generalizzato, Fig.(1a), il tipo di utensile, la lunghezza, il diametro l'inclinazione dell'elica del tagliente ed il numero di taglienti possono essere modificati, in modo da poter simulare le forze di taglio, per diverse tipologie di utensile, Fig.(1b) e Fig.(1c).

La previsione delle forze di taglio, necessita della definizione dei 6 coefficienti di taglio, tali grandezze sono delle costanti per date condizioni di taglio: parametri di processo, geometria dell'utensile e materiali pezzo ed utensile.

Le componenti delle forze di taglio in Eq.(3), sono scritte rispetto ad un sistema di riferimento locale, esse sono poi trasformate secondo un sistema di riferimento assoluto (cartesiano), così come in Eq.(7).

$$\begin{bmatrix} dF_x \\ dF_y \\ dF_z \end{bmatrix} = \begin{bmatrix} -\sin \phi_j \cdot \sin \kappa & -\cos \phi_j & -\sin \phi_j \cdot \cos \kappa \\ -\cos \phi_j \cdot \sin \kappa & \sin \phi_j & -\cos \phi_j \cdot \cos \kappa \\ -\cos \kappa & 0 & -\sin \kappa \end{bmatrix} \cdot \begin{bmatrix} dF_r \\ dF_t \\ dF_a \end{bmatrix} \quad (7)$$

Le forze di taglio si ottengono dalla integrazione delle Eq.(7) lungo la profondità di passata (all'interno di z_1, z_2) in Eq.(8).

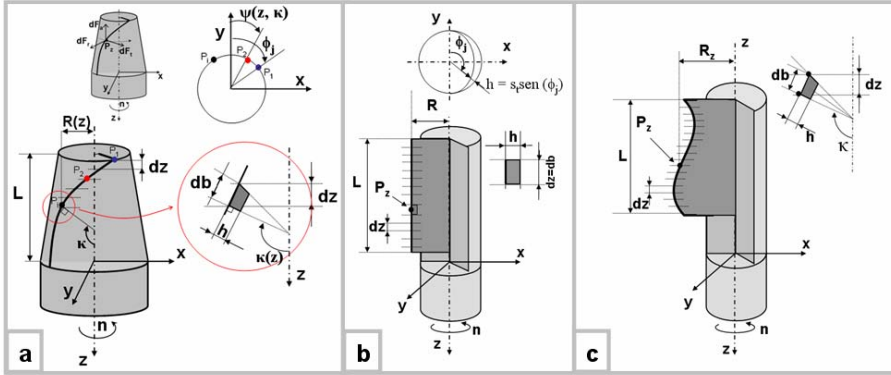


Figura 1 a) Modello generalizzato dell'utensile, b) e c) differenti tipologie di parametrizzazioni.

$$\begin{aligned}
 F_x &= \sum_{j=1}^N \int_{z_1}^{z_2} \left[-dF_{rj} \cdot \sin \phi_j \cdot \sin k_j - dF_{tj} \cdot \cos \phi_j - dF_{aj} \cdot \sin \phi_j \cdot \cos k_j \right] \cdot dz \\
 F_y &= \sum_{j=1}^N \int_{z_1}^{z_2} \left[-dF_{rj} \cdot \cos \phi_j \cdot \sin k_j + dF_{tj} \cdot \sin \phi_j - dF_{aj} \cdot \cos \phi_j \cdot \cos k_j \right] \cdot dz \\
 F_z &= \sum_{j=1}^N \int_{z_1}^{z_2} \left[-dF_{rj} \cdot \cos k_j - dF_{aj} \cdot \sin k_j \right] \cdot dz
 \end{aligned} \tag{8}$$

La profondità di taglio è suddivisa in un numero di incrementi Δz lungo l'asse Z (asse dell'utensile). Quindi, per ogni incremento Δz , si hanno i parametri corrispondenti (dS , h , db) che vengono calcolati.

L' Eq.(8) è risolta ad ogni passo di integrazione, quando il punto P_z è ingaggiato con il pezzo in lavorazione (quindi per $\phi_{st} < \phi_{ex}$), e quando h diventa maggiore di zero.

La definizione degli angoli di inizio e fine ingaggio (ϕ_{sb} , ϕ_{ex}) permette di realizzare il calcolo delle forze di taglio, per diverse condizioni di taglio stesso come down-mill, up-mill and slot-mill come mostrato in Fig.(2).

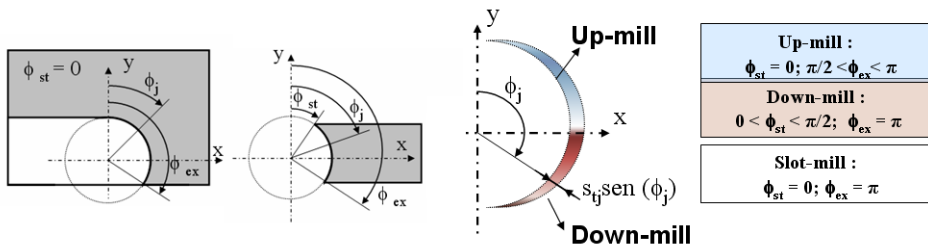


Figura2. Parametri: ϕ_{st} e ϕ_{ex} .

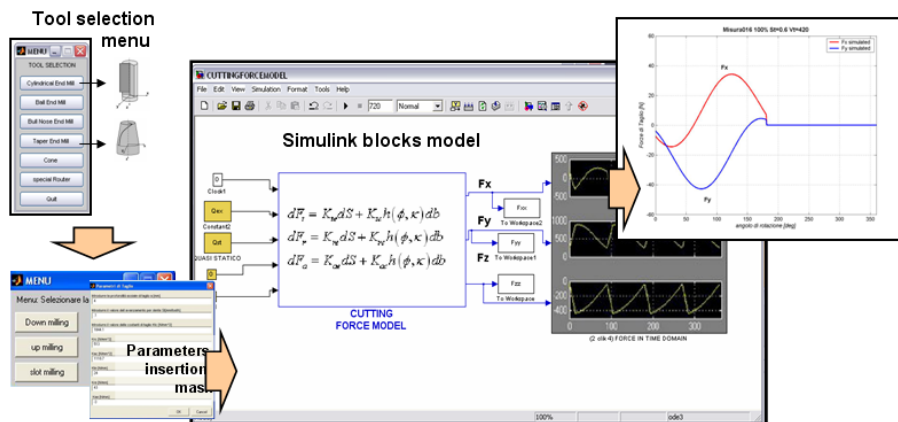


Figura 3: Modello implementato in Simulink.

Lo studio della soluzione presentata è organizzato in moduli, (Fig.(3)) tali moduli sono assemblati l'uno all'altro una parte di tali blocchi sono adibiti all'inserimento dei parametri (forma e dimensioni utensile) altre sono dedicate al calcolo a altre ancora alla redazione dei risultati.

È quindi possibile calcolare la coppia che agisce sul mandrino e la Potenza necessaria per effettuare il taglio.

3. COEFFICIENTI DELLE FORZE DI TAGLIO

Una campagna sperimentale è stata realizzata al fine di ottenere la stima dei coefficienti K secondo le procedure in [10,11].

La forza di taglio media di taglio, per rotazione dell'utensile è riportata grazie alle Eq(9).

$$\begin{cases} \overline{F_x} = -K_{te}S + K_{re}T - \frac{S_t}{4}(-K_{tc}P + K_{rc}Q) \\ \overline{F_y} = -K_{te}T - K_{re}S + \frac{S_t}{4}(K_{tc}Q + K_{rc}P) \\ \overline{F_z} = -\frac{aN}{2\pi}K_{ae}(\vartheta_{ex} - \vartheta_{st}) + S_t K_{ac}T \end{cases} \quad (9)$$

$$P = \frac{aN}{2\pi}[\cos 2\vartheta]_{\vartheta_{st}}^{\vartheta_{ex}}; \quad S = \frac{aN}{2\pi}[\sin \vartheta]_{\vartheta_{st}}^{\vartheta_{ex}}$$

$$Q = \frac{aN}{2\pi}[2\vartheta - \sin 2\vartheta]_{\vartheta_{st}}^{\vartheta_{ex}} \quad T = \frac{aN}{2\pi}[\cos \vartheta]_{\vartheta_{st}}^{\vartheta_{ex}} \quad (10)$$

Dove $a = z_2 - z_1$. ed i parametri P, Q, T, S sono delle costanti durante l'esecuzione della prova, e sono calcolati tramite le Eq.(10).

Ipotizzando una dipendenza lineare tra le forze di taglio e l'avanzamento per dente è possibile scrivere le Eq.(11), dove la forza di taglio ha due componenti principali: forze di bordo (F_e) e componenti delle forze di taglio (F_c).

$$\begin{cases} \overline{F_x} = \overline{F_{xe}} + s_t \overline{F_{xc}} \\ \overline{F_y} = \overline{F_{ye}} + s_t \overline{F_{yc}} \\ \overline{F_z} = \overline{F_{ze}} + s_t \overline{F_{zc}} \end{cases} \quad (11)$$

Nelle Eq.(11) le forze (sopra segnate) sono misurate durante la prova e I coefficienti di taglio possono essere ricavati grazie alle Eq(12).

$$\begin{aligned} K_{ic} &= 4 \frac{\overline{F_{xc}} P + \overline{F_{yc}} Q}{P^2 + Q^2} & K_{ac} &= \frac{\overline{F_{zc}}}{T} & K_{re} &= \frac{K_{ie} S + \overline{F_{xe}}}{T} \\ K_{rc} &= \frac{K_{ic} P - 4 \overline{F_{xc}}}{Q} & K_{te} &= -\frac{\overline{F_{xe}} S + \overline{F_{ye}} T}{S^2 + T^2} & K_{ae} &= -\frac{2\pi}{aN} \frac{\overline{F_{ze}}}{(\phi_{ex} - \phi_{st})} \end{aligned} \quad (12)$$

4. PROVE SPERIMENTALI

Pannelli di MDF, di spessore di 19 mm e umidità specifica del 12% sono fresati con un utensile cilindrico di 10 mm di raggio, l'utensile presenta un tagliente dritto, così come mostrato in Fig.(4), ed è composto da metallo sinterizzato la macchina utilizzata è un 5 assi CNC: SCM Record 110 AL Prisma, Fig.(4).

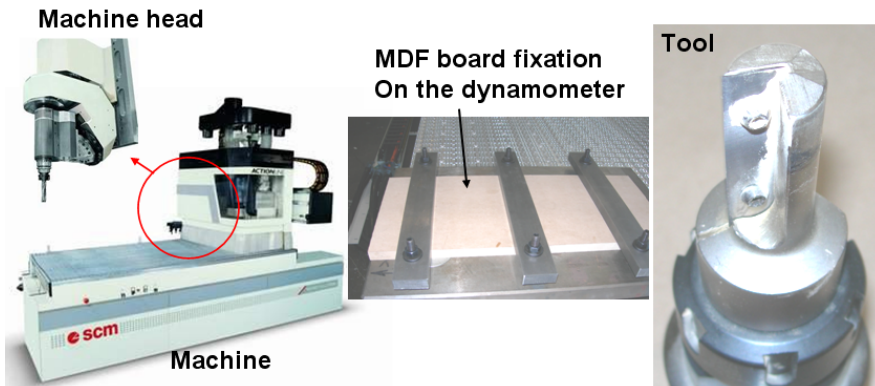


Figura 4. Macchina 5 assi controllati utilizzata; utensile cilindrico tavola dinamometrica, con pannello fissato;

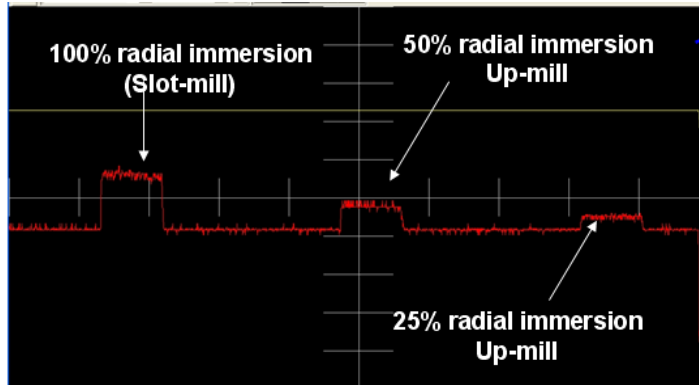


Figura 5: Segnale acquisito da oscilloscopio digitale.

Le forze sono registrate con frequenza di acquisizione 12 kHz e successivamente filtrate grazie a filtri passa basso e passa alto di tipo Butterworth. L'ordine dei filtri applicati e le frequenze di taglio sono differenti per ogni segnale.

I coefficienti delle forze di taglio sono determinate grazie alle Eq.(12).

Durante i test di lavorazione, sono state acquisiti anche gli andamenti delle potenze di taglio, tali grandezze sono state misurate grazie ad un oscilloscopio digitale collegato all'inverter del mandrino della macchina, Fig.(5) . grazie a queste misurazioni è possibile valutare secondo un'altra strada le forze di taglio medie, utilizzando l'Eq.(13) riportata in Fig.(5).

$$\begin{aligned}
 P_t &= \overline{M_t} \cdot n[\text{rpm}] \cdot \frac{2\pi}{60} \\
 \overline{M_t} \cdot 2\pi &= M_t (\phi_{ex} - \phi_{st}) \\
 F_t &= \frac{60}{n} \cdot \frac{1}{z \cdot \varphi} \cdot \frac{1000}{R}
 \end{aligned} \tag{13}$$

5. RISULTATI, SIMULAZIONI E CONFRONTI

L'analisi statistica del piano fattoriale utilizzato per le prove sperimentali conferma un buon grado di significatività dei fattori, mentre la distribuzione dei residui assicura una bassa varianza dei dati. L'analisi della varianza conferma inoltre che la principale assunzione fatta (in Eq.(11)) è ben posta, ciò è molto importante da verificare al fine della attendibilità dei valori dei coefficienti di taglio calcolati. In Tabella 1 sono riportati i valori delle costanti K . per il caso di 100% Slot mill calcolati grazie alle Eq.(12).

Una volta che le costanti K sono determinate è possibile condurre delle simulazioni di forza di taglio, inserendo gli le stesse, le condizioni operative e le informazioni geometriche dell'utensile nel foglio di lavoro in Simulink.

Alcuni confronti tra simulazioni e misura sono riportate in Fig.(6). le forze di taglio sono plottate come funzioni dell'angolo di rotazione istantaneo, ma nella assunzione di una modellazione quasi stazionaria, possono essere riferite al tempo.

Tabella 1. Costanti K per 100% Slot mill.

K.. coefficients for 100% slot mill operation	Vt = 420 [m/min]	Vt = 400 [m/min]	Vt = 380 [m/min]
	Ktc = 5,7002	Ktc = 6.0442	Ktc = 3.9425
Krc = -3,0374	Krc = -2.2866	Krc = -1.8781	
Kac, Kae =0 (Cylindrical tool shape)	Kte = 0,7807	Kte = 0.4609	Kte = 0.4367
	Kre = -0,3995	Kre = -0.5463	Kre = -0.5574

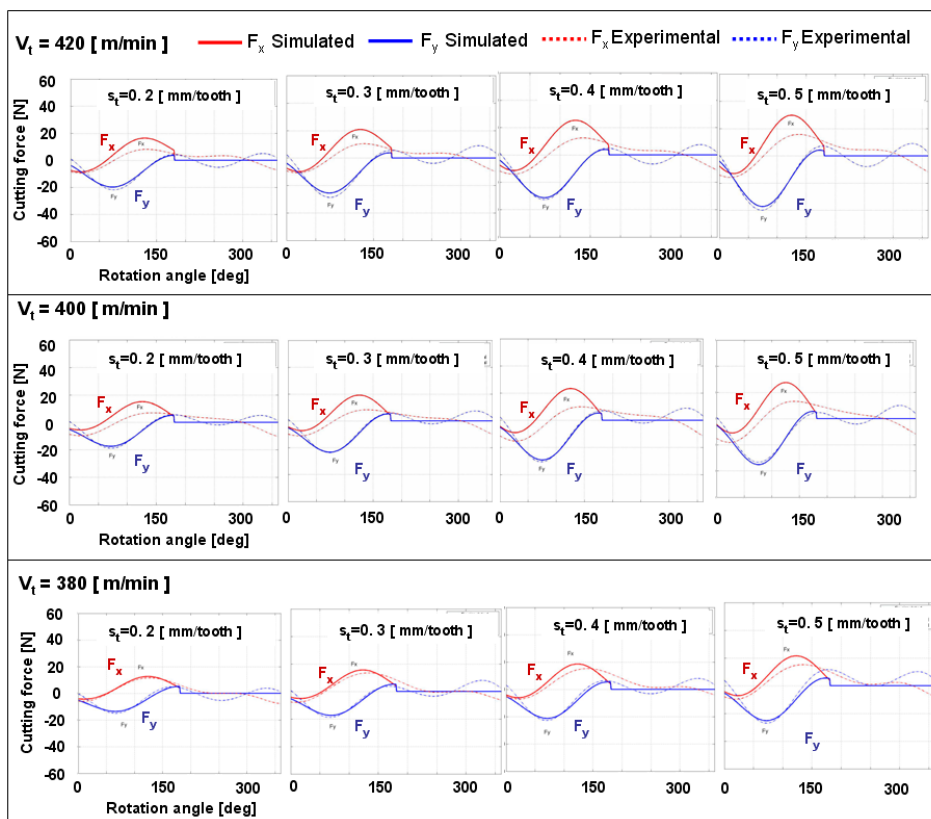


Figura 6: confronto tra simulazioni e misure sperimentali, nel caso di lavorazioni in slot-mill.

Tabella 2: confronto delle forze medie di taglio ($v_i=420\text{m/mm}$, Slot-mill).

Feed per tooth	Measured mean cutting force (from the piezoelectrics dynamometer)	Calculated mean cutting force (from Simulink implemented model)	Error
s_t [mm/dente]	$F_{t_dynamometer}$ [N]	$F_{t_simulated}$ [N]	[%]
0.2	5.66	7.27	28
0.3	7.45	9.03	21
0.4	8.8	10.8	23
0.5	10	12.6	25
0.6	12	14.4	19

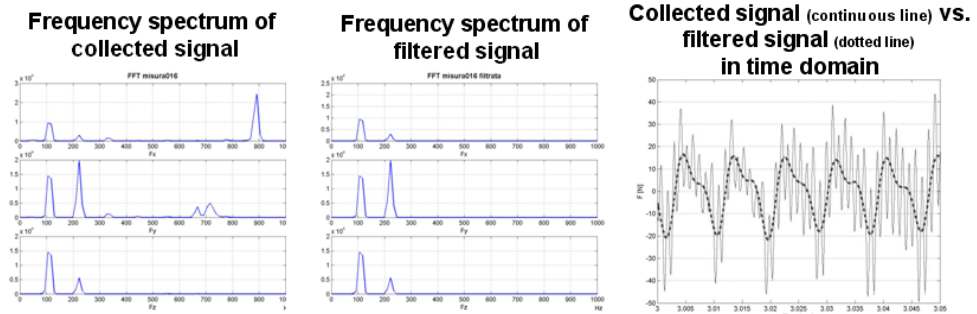


Figura 7: segnale acquisito e segnale filtrato

La forza media di taglio, calcolata e agente durante il periodo di ingaggio di un dente $F_{t_simulated} = \sqrt{F_x^2 + F_y^2}$ è stata anch'essa valutata. la forza media misurata tramite dinamometro è invece esprimibile come: $F_{t_dynamometer} = \sqrt{\overline{F_x^2} + \overline{F_y^2}}$.

In Tabella 2 sono mostrati i valori delle forze medie, misurate e calcolate, con il rispettivo errore.

I valori degli errori sono abbastanza alti, ciò è dovuto alla eccitazione del primo modo di vibrare della tavola dinamometrica, da parte della forzante.

Come mostrato in Fig.(7) il segnale filtrato è affetto dalla vibrazione della tavola dinamometrica.

Una ulteriore indagine è stata effettuata partendo dalle potenze di taglio acquisite dall'inverter, al fine di determinare il k_c -value descritto dal modello di Kivimaa/ Pahlitzsch [2, 3].

È stato quindi valutato il valore del k_c -value, così come da Eq.(1). dove b è la larghezza di taglio e h_m è lo spessore di truciolo medio che viene valutato (nel caso di geometria dell'utensile, come nel nostro caso) dalla Eq.(14), per il caso di slot mill, mentre per operazione di 25% Up-mill ci si riferisca all'Eq.(15).

Applicando l'Eq(1), la forza media di taglio F_c viene calcolata per diversi valori di s_t , considerando l'equazione $k_c = w_0 + \frac{K}{h_m}$, e applicando una regressione lineare si ricava il k_c -values.

$$h_m \cdot R \cdot (\phi_{ex} - \phi_{st}) = A_{chip} \cong 2R \cdot s_t \Rightarrow h_m = \frac{2 \cdot s_t}{(\phi_{ex} - \phi_{st})} \quad (14)$$

$$h_m \cdot R \cdot (\phi_{ex} - \phi_{st}) = A_{chip} \cong R/2 \cdot s_t \Rightarrow h_m = \frac{s_t}{2(\phi_{ex} - \phi_{st})} \quad (15)$$

In tabella 3 sono ripostati i valori medi delle potenze di taglio acquisite dall'inverter del mandrino, la forza media di taglio è fornita da Eq.13 e il k_c -values valutato attraverso regressione lineare di questi dati. Sempre in tabella 3 sono inoltre confrontati i della forza media di taglio calcolati e misurati in questa procedura.

In tabella 4 sono comparati I risultati di tali coefficienti trovati in letteratura, con quelli calcolati in questo lavoro

La differenza dei valori di tali parametri è sostanzialmente giustificata dalle diverse, macchine, utensili e condizioni operative del caso, e in maggior parte dalla diversa natura e provenienza della fibra di legno costituente i pannelli .

Infine un ultimo confronto è riportato in tabella 5 dove sono confrontati I valori mediate delle forze di taglio, ricavati dalle 2 procedure, sia nel caso di calcolo, che nel caso di misura.

6. CONCLUSIONI.

In questo lavoro sono stati comparati diversi modelli per la previsione della forze di taglio agenti, nella fresatura di pannelli di MDF. I modelli proposti sono stati implementati in ambiente software al fine di fornire uno strumento di calcolo utile a livello aziendale.

Si è dimostrato che è possibile trasferire modellazioni usualmente utilizzate nel taglio di metalli, anche in alcuni particolari casi di taglio di materiali compositi, in questo caso a fibra legnosa.

Le forze di taglio vengono espresse come funzioni dell'avanzamento per dente, delle condizioni di ingaggio pezzo utensile, e dalle condizioni di taglio. I coefficienti di taglio (K) possono essere ricavati da prove di lavorazione su macchina strumentata.

Grazie a questi modelli è possibile predire la distribuzione delle componenti delle forze di taglio rispetto al tempo (o all'angolo di rotazione).

Il modello di Kivimaa/ Pahlitzsch è un modello meno complesso, che richiede meno strumentazione e know how, ma che fornisce minori caratteristiche.

Tabella 3: dati acquisiti e calcolati a partire dalle potenze di taglio assorbite.

Feed per tooth	FROM MEASURED CUTTING POWER		CALCULATED BY THE AUTHORS		Error
	P_t [W]	F_t [N]	k_c -values [N/mm ²]	F_t [N]	
s_t [mm/tooth]					
0.2	18.06	5.16	4.40	5.32	-3.12
0.3	25.98	7.42	3.81	6.91	6.86
0.4	29.75	8.50	3.52	8.51	-0.07
0.5	34.63	9.89	3.34	10.10	-2.06
0.6	39.93	11.41	3.22	11.69	-2.47

Tabella 4: K and w parameters comparison, literature data taken from [4]

Calculated by	AUTHORS	Tröger	Kivimaa	Scholz	IHF Dresden
K [Wsmm/cm ³]	0.3	2.4	1.6	1.5	1.0
w ₀ [Ws/cm ³]	2.6	7.0	8.0	48	13.0

Tabella 5: Confronto finale.

Feed per tooth	Measured	Measured	Calculated Follow Kivimaa	Calculated Follow Altintas
S _t [mm/tooth]	F _t dynamometer [N]	F _t Inverter [N]	F _t calculated [N]	F _t simulated [N]
0.2	5.66	5.16	5.32	7.27
0.3	7.45	7.42	6.91	9.03
0.4	8.8	8.50	8.51	10.8
0.5	10	9.89	10.10	12.6
0.6	12	11.41	11.69	14.4

BIBLIOGRAFIA

- [1] A. Fortunato, C. Mantega, L. Donati, G. Tani. 2003, Milling force prediction by means of analytical model and 3D FEM simulations, Proceedings of the 9th CIRP International Workshop on Modeling of Machining Operations.
- [2] E. Kivimaa. 1950 Cutting force in wood-work, PhD thesis, Finland State Inst. For tech, Research, Helsinki, Finland.
- [3] G. Pahlitzsch, H. Jostmeier. 1956, Untersuchungen beim Frasen von Spanplatten und Schichtsoff-Verbundplatten. Moderne Holzverarbeitung p 66.
- [4] F. Scholz, J. Troeger, 2005, Modelling of Cutting Forces, in: Proceedings of the 17th International Wood Machining Seminar part2: Posters , pp. 1-5.
- [5] U. Heisel, J. Troeger, H. Dietz, 1995, Am Schneidkeil wirkende Krafte (2), *Mitteilung HK*, **6**, pp.884-888.
- [6] W.M. McKenzie.1956 Fundamental aspects of wood cutting process, *Forest Production Journal*, **41** (9), pp. 61-64.
- [7] Y. Altintas, P. Lee; 1996, Prediction of ball-end milling forces from orthogonal cutting data; *International Journal of Machine Tools and Manufacture*, **36** (9) , pp. 1059-1072.
- [8] Y. Altintas, S. Engin, 2001. Generalized modeling of milling mechanics and Dynamics: part I- Helical End Mills, *International Journal of Machine Tools and Manufacture*, **41**(15), pp. 2195-2212.
- [9] S. Engin, Y. Altintas, F.B. Amara. 2000 Mechanics of routing Medium Density Fiberboard, in: *Forest Products Journal*, **50** (9), pp. 65-69.
- [10] E. Budak, Y. Altintas, E. J. A. Armarego. 1996, Prediction of Milling Force Coefficients From Orthogonal Cutting Data, *Transactions of ASME, Journal of Manufacturing Science and Engineering*, **118**, pp. 216-224 .
- [11] J. Gradisek, M.Kalveram, K. Weinert. 2004 , Mechanistic identification of specific force coefficients for a general end mill; *International Journal of Machine Tools and Manufacture*, **44**, pp . 401-414.

[12] J.P. Costes, C. Decès-Petit, Y. Altintas, P. Ko. 1986, Estimated stress and friction distribution on tool rake in the medium density fiberboard cutting process; *Forest Products Journal*, **53**, pp. 59-66.

MODELLAZIONE DEL PROCESSO DI TRATTAMENTO TERMICO SUPERFICIALE MEDIANTE FASCIO LASER

Leonardo Orazi

Department of Science and Methods for Engineering,

University of Modena and Reggio Emilia, Italy

E-mail: leonardo.orazi@unimore.it

Alessandro Fortunato

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: alessandro.fortunato@unibo.it

Giovanni Tani

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: giovanni.tani2@unibo.it

Gabriele Cuccolini

Department of Science and Methods for Engineering,

University of Modena and Reggio Emilia, Italy

E-mail: gabriele.cuccolini@unimore.it

Alessandro Ascari

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: alessandro.ascari@unibo.it

Giampaolo Campana

Department of Mechanical and Aeronautical Engineering,

University of Bologna, Italy

E-mail: giampaolo.campana@unibo.it

Abstract. *Il lavoro presentato in quest'articolo mette in evidenza i vantaggi e le problematiche relative al processo di tempra superficiale laser relativamente nuovo ma in forte espansione nelle aziende manifatturiere. In particolare, l'articolo si concentra sulle problematiche tipiche di questa tecnologia e propone, in tale ambito, le soluzioni adottate per far fronte a tali problematiche ed un approccio volto all'ottimizzazione dei principali parametri in gioco mediante l'utilizzo di un software di simulazione di processo.*

Keywords: *laser, trattamenti superficiali, tempra, rinvenimento*

1. INTRODUZIONE

La tempra laser è un processo che ha il fine di indurire, mediante trattamento termico, superfici di componenti realizzati in acciaio a medio contenuto di carbonio. Gli spessori induriti, in relazione alla quantità di calore trasferito al materiale, variano da qualche centinaio di micron al millimetro con durezza superiori ai 700 HV.

Tale processo di indurimento superficiale negli ultimi anni sta riscontrando un forte interesse nelle aziende manifatturiere dal momento che presenta indubbi vantaggi se paragonato alle tecnologie di indurimento superficiali tradizionali come la tempra a fiamma o ad induzione, vantaggi riassumibili in:

- "pulizia di processo": la tempra laser non necessita infatti di mezzi raffreddanti per ottenere elevate durezza;
- processo estremamente "selettivo" nelle geometrie trattabili: si possono indurire zone limitate del componente, anche molto piccole al limite delle dimensioni dello spot del laser, consentendo, quindi, di trattare solo le parti che devono essere effettivamente indurite.

Inoltre, scegliendo la sorgente laser opportuna, il trasferimento dell'energia avviene per assorbimento della radiazione stessa e non per dissipazione di correnti indotte aumentando notevolmente l'efficienza energetica del processo. A tale proposito vi è da tenere in considerazione che normali induttori elettromagnetici per tempra superficiale hanno potenze installate dell'ordine dei $50 \div 100 \text{ kW}$ a fronte di sorgenti laser con potenze nominali dell'ordine di $1 \div 2 \text{ kW}$. Da quanto detto, al vantaggio economico della riduzione energetica, si aggiunge un ulteriore vantaggio tecnologico dovuto al notevole contenimento delle deformazioni, e quindi delle tensioni, indotte dal processo sul pezzo dopo il trattamento. Nonostante i vantaggi elencati in precedenza alcune problematiche tecnologiche devono ancora essere risolte. La prima, e forse più importante su cui questo lavoro si concentra, riguarda la scelta della strategia ottimale della modalità di scansione della superficie da trattare da parte del raggio poiché, quando si devono trattare superfici più grandi dello spot del raggio laser, sono necessarie passate multiple le quali a loro volta rinvengono le zone precedentemente indurite [1]. La scelta della strategia di scansione ottimale, in termini di velocità di scansione, scelta del grado di sovrapposizione fra passate contigue e della tipologia di scansione, possono incrementare di molto la resa del processo e quindi aumentare la sua applicabilità in campo industriale su pezzi anche di grandi dimensioni. Come tutti i processi dipendenti dalla temperatura, diventa fondamentale conoscere come tale temperatura varia nel componente da trattare dal momento che la condizione necessaria per far avvenire il trattamento termico è che si raggiunga, per ciascun punto interessato al trattamento, la temperatura di austenitizzazione e che tale temperatura venga mantenuta per il tempo necessario a fare avvenire la dissoluzione della perlite [2, 3]. Un'ulteriore complicazione concettuale tipica del trattamento termico laser è che l'austenitizzazione avviene a temperature maggiori, anche di centinaia di gradi, di quelle previste nei normali diagrammi di stato poiché i cicli termici sono estremamente rapidi e l'austenitizzazione, che è un fenomeno legato alla diffusione del carbonio, non può che avvenire a temperature maggiori.

Infine, trattandosi di un trattamento termico, diventa fondamentale non solo conoscere la modalità di riscaldamento del componente ma anche la velocità con cui esso si raffredda dal momento che è l'assorbimento di calore verso l'interno del materiale che deve garantire il superamento della velocità critica di tempra poiché, come già detto, la tempra laser

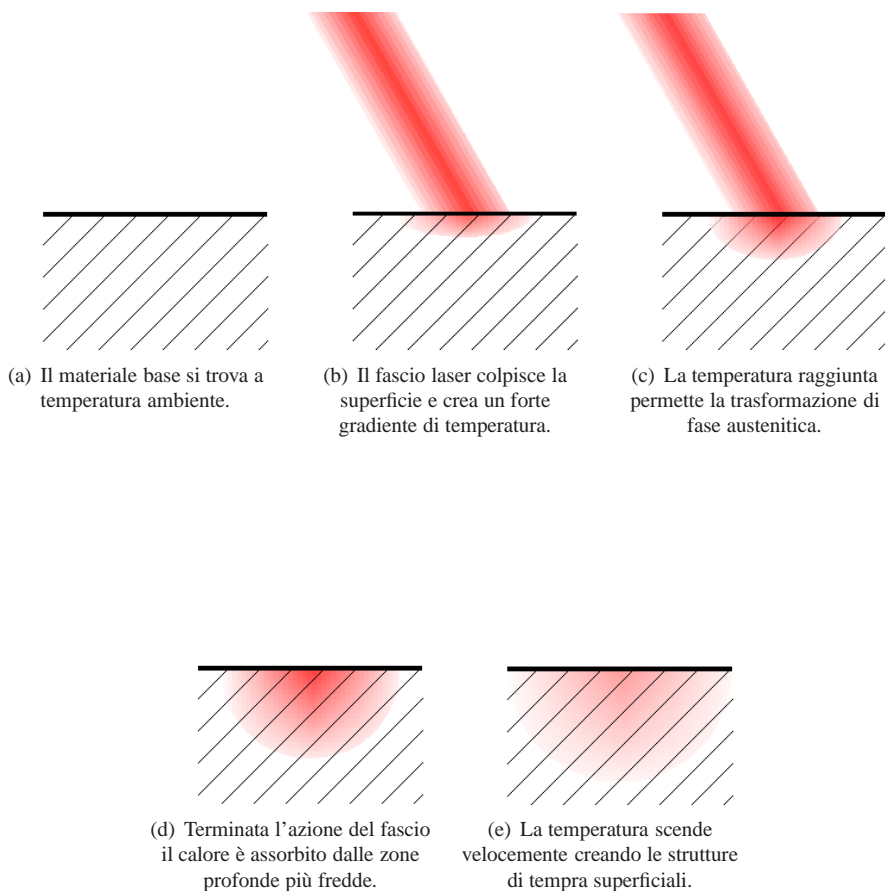


Figure 1: Schematizzazione del processo di tempra laser superficiale di un componente massivo.

non richiede mezzi raffreddanti esterni. La figura 1 descrive schematicamente il processo mettendo in evidenza i punti salienti nell'interazione laser-materiale.

Il complesso dei fenomeni legati all'assorbimento della radiazione laser, funzione del materiale da trattare e della sua finitura superficiale, al ciclo termico, inteso come riscaldamento e raffreddamento in funzione del tipo di materiale e della geometria del componente, alla diffusione degli elementi di lega [4], alla scelta del percorso ottimale, inteso come velocità e strategia di scansione, alle microstrutture risultanti [5–7] e alle durezze ottenute dopo il trattamento costituiscono problematiche di difficile ottimizzazione sperimentale e per questo lo sviluppo di un simulatore diventa di estremo aiuto alla messa a punto ottimale del processo.

Tale simulatore denominato Laser Hardening Simulator (LHS) verrà presentato di seguito insieme alle problematiche principali che si incontrano nella modellazione del processo della tempra superficiale laser.

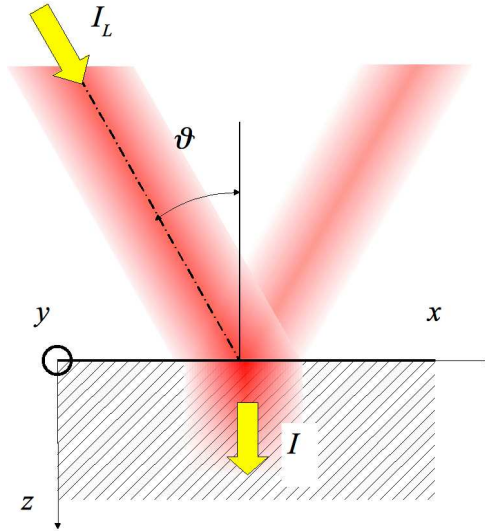


Figure 2: Il software LHS tiene conto di una qualsiasi distribuzione del fascio, inclinazione, riflettività ed assorbimento del materiale.

2. IL SOFTWARE “LHS”

Il software di simulazione LHS presentato in questa sede consente, mediante il metodo alle differenze finite, di calcolare punto per punto ed in condizioni non stazionarie il ciclo termico a cui è soggetto il componente meccanico, nonché la conseguente diffusione del carbonio all’interno della struttura cristallina. Di particolare interesse risulta l’algoritmo per determinare la potenza termica assorbita dal materiale, figura 2, valutata tenendo conto di una qualsiasi distribuzione di intensità del fascio laser e di condizione di focalizzazione dello stesso nonché dei coefficienti di riflessione e di assorbimento della radiazione elettromagnetica da parte del materiale. Questi parametri possono essere fatti variare in funzione della temperatura come normalmente accade e permettono di tenere conto delle differenti condizioni superficiali dei particolari da trattare. A fronte di questi risultati si è quindi in grado di modellare i fenomeni di dissoluzione della perlite, di omogeneizzazione del carbonio all’interno dei grani austenitici e di formazione, durante il raffreddamento, delle conseguenti strutture martensitiche, bainitiche, ferritiche e perlitiche.

I risultati in uscita riguardano dunque la previsione non solo della profondità e della durezza dello strato temprato, ma anche della relativa composizione in termini di fasi metalliche e dell’eventuale effetto di rinvenimento dovuto al mutuo effetto di passate contigue. Mediante l’assegnazione di opportuni “probes” l’utente è in grado di estrapolare tutti questi risultati previsionali in punti precisi del pezzo in lavorazione.

Il software è strutturato in modo da acquisire in ingresso le informazioni riguardanti la geometria del pezzo, la struttura metallografica ed i parametri fisici del materiale. La discretizzazione può essere effettuata direttamente da LHS se la geometria è semplice, mentre per i componenti più complessi si avvale di CUBIT, un pre-processor esterno avanzato. I parametri laser sono inseriti in termini di densità di potenza, distribuzione di potenza e di-

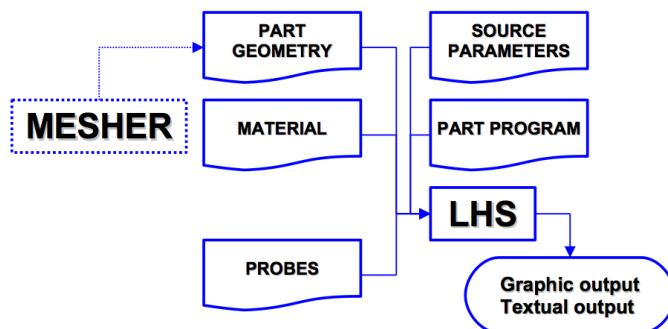


Figure 3: Schematizzazione delle operazioni in LHS del processo di tempra laser.

ametro dello spot di lavorazione, mentre lo specifico percorso di lavorazione è impostato mediante linguaggio ISO-base. Tra le caratteristiche maggiormente apprezzabili del sistema in oggetto spicca la possibilità di ottenere in uscita il percorso ISO ottimizzato e i relativi parametri laser ottimali in caso di lavorazioni complesse in cui si richieda di minimizzare gli effetti di rinvenimento dovuti a passate ravvicinate. In tal senso si può quindi parlare di un vero e proprio CAM per lavorazioni di trattamento superficiale mediante sorgente laser.

Uno schema logico delle operazioni di simulazione di processo mediante il software LHS è mostrato in figura 3.

3. GLI ALGORITMI PER L'AUSTENITIZZAZIONE RAPIDA

I principi fisici e le leggi che governano l'austenitizzazione delle leghe ferrose in condizioni prossime all'equilibrio termodinamico sono ben noti grazie alla notevole quantità di studi condotta su questo argomento. Nei trattamenti termici tradizionali il materiale viene posto in forno e portato a temperature tra i 50 ed i 90 °C sopra la temperatura A_{c3} quindi mantenuto a queste temperature per un tempo dipendente dallo spessore (indicativamente qualche minuto per ogni millimetro di spessore). Lo scopo è di ottenere una struttura austenitica con il carbonio distribuito uniformemente nel reticolo cristallino evitando, al contempo, l'eccessivo ingrossamento dei grani cristallini.

Ben diversa è la situazione per quanto riguarda processi e trasformazioni in condizioni lontane dall'equilibrio come nel caso della tempra laser superficiale. In questo caso si può supporre che la trasformazione austenitica possa essere divisa in due fasi:

- Diffusione intragranulare del carbonio all'interno delle strutture perlitiche.
- Diffusione intergranulare del carbonio tra i grani perlitici e ferritici.

Al termine della prima fase si può supporre la struttura costituita da una miscela di grani austenitici con composizione eutettica e di grani precedentemente ferritici con tenore di carbonio praticamente nullo. Con il procedere del riscaldamento il carbonio diffonde su distanze dell'ordine della dimensione dei grani portando alla formazione di austenite omogenea.

I primi modelli relativi alle trasformazioni rapide sono stati proposti in [2, 8, 9]. In particolare in [2] si afferma che nelle trasformazioni rapide la diffusione del carbonio all'interno

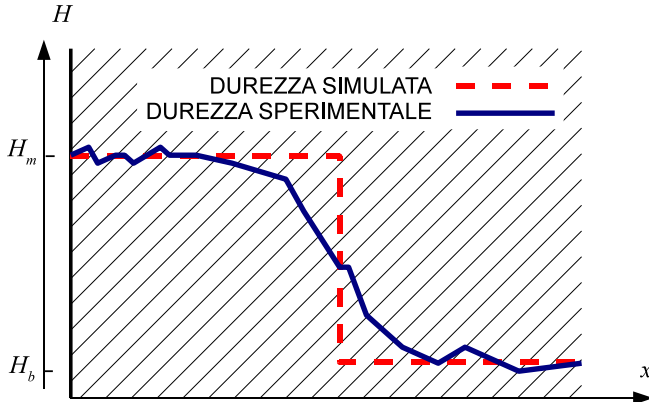


Figure 4: Confronto tra un tipico profilo di durezza sperimentale e uno simulato.

del grano perlitico avvenga in direzione normale alle lamelle di cementite ma la mancata corrispondenza tra i valori del surriscaldamento misurati e quelli predetti portano poi gli autori a supporre che la diffusione avvenga solo attraverso l'estremità delle lamelle secondo l'equazione 1 dove D_0 è la diffusività del carbonio, Q l'energia di attivazione del processo, R la costante dei gas perfetti, $T(t)$ l'andamento della temperatura nel tempo, l'integrale è poi calcolato tra gli estremi $t_{A_{c1}}$ e $t_{A_{r1}}$, tempi ai quali la temperatura rimane al di sopra della temperatura A_1 , il parametro di confronto è dato dalle dimensioni trasversali della lamella caratterizzata dalla dimensione della colonia perlitica L e dallo spessore della lamella di cementite λ .

$$\int_{t_{A_{c1}}}^{t_{A_{r1}}} D_0 \exp\left(-\frac{Q}{RT(t)}\right) dt \geq L\lambda \quad (1)$$

Un modello molto accurato per la formazione di austenite a partire da perlite lamellare in caso di trattamenti in forno è proposto in [4] e successivamente dettagliato in [10] e [11]. In particolare gli autori evidenziano come la diffusione del carbonio, che in condizioni di quasi equilibrio avviene attraverso l'estremità della lamella, diventi per velocità di riscaldamento di poche decine di gradi al secondo normale alla lamella stessa. Ciò indica come il modello descritto dall'equazione 1 e tutti quelli da esso derivati siano fortemente limitati, dato che prendono in esame una configurazione di diffusione del carbonio adatta per situazioni di equilibrio. Un'altra criticità dei modelli presenti è che non tengono normalmente conto della zona di transizione delle durezza nello spessore trattato. La figura 4 mostra il confronto tra un andamento della durezza nello spessore misurato ed il risultato di un modello basato sulla equazione 1, nell'immagine H_m è la durezza della struttura martensitica ottenuta mentre H_b la durezza del materiale base. Si nota che il comportamento a gradino delle durezza simulate contrasta con la presenza di una zona di transizione in molti casi di estensione analoga o maggiore della zona a massima durezza.

Il comportamento a gradino delle durezza simulate deriva dal fatto che i parametri L e λ sono considerate costanti in tutti i punti del materiale, cosa ovviamente non corretta essendo parametri microstrutturali uniformemente distribuiti.

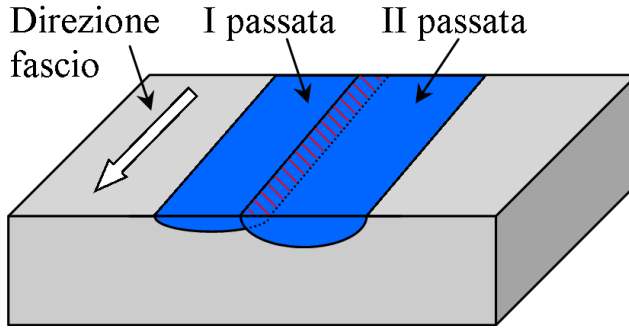


Figure 5: Il problema del sormonto.

Per ovviare a questo problema in questo lavoro si propone un modello che considera un parametro microstrutturale non costante nel materiale che tenga conto dei problemi esposti e sia efficacemente applicabile dal punto di vista computazionale. Il parametro, definito dall'equazione 2 è denominato $I_{p \rightarrow a}$ e rappresenta dimensionalmente il tempo richiesto per la trasformazione da struttura perlitica ad austenitica.

$$I_{p \rightarrow a} = \int_{t_{Ac1}}^{t_{Ar1}} \exp\left(-\frac{Q_{p \rightarrow a}}{RT(t)}\right) dt \quad (2)$$

Il modello presuppone che la trasformazione abbia luogo quando il valore di $I_{p \rightarrow a}$ supera una soglia come da equazione 3.

$$I_{p \rightarrow a} \geq I_{p \rightarrow a, th} \quad (3)$$

La formazione di una zona di transizione può essere modellata supponendo che la variabile di soglia $I_{p \rightarrow a, th}$ sia statisticamente distribuita all'interno del materiale così come risultano distribuite le caratteristiche microstrutturali (dimensioni dei grani, spessore delle lamelle di perlite, ecc...). Tra tutte le possibili distribuzioni statistiche viene scelta, per ragioni computazionali la distribuzione uniforme caratterizzata dai valori estremi di esistenza $I_{p \rightarrow a, min}$ e $I_{p \rightarrow a, max}$. Sotto questa ipotesi la frazione di austenite che si genera, a partire dalla perlite esistente è ottenibile dalla equazione 4.

$$f_a = \frac{I_{p \rightarrow a} - I_{p \rightarrow a, min}}{I_{p \rightarrow a, max} - I_{p \rightarrow a, min}} \quad (4)$$

Le elevate velocità di raffreddamento tipiche di questo tipo di processo garantiscono che praticamente tutta l'austenite presente si trasformi in martensite quindi la frazione di martensite f_m generata nel raffreddamento risulta, punto per punto, pari a f_a , noto questo valore è possibile stimare il valore di durezza utilizzando le usuali formule reperite in letteratura.

4. IL PROBLEMA DEL RINVENIMENTO

Uno dei limiti del processo di tempra laser è la limitata dimensione delle zone che possono essere trattate con una singola passata. La presenza di estese superfici da trattare o di percorsi di trattamento "chiusi" costringe a realizzare passate adiacenti una all'altra con un certo grado

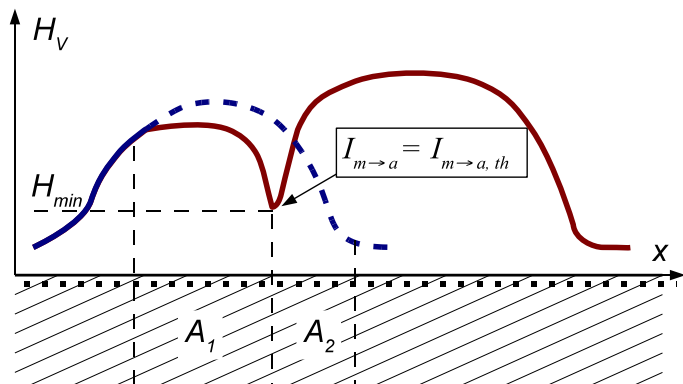


Figure 6: Andamento della durezza in presenza di due passate sovrapposte.

di sovrapposizione. Gli effetti dovuti alla sovrapposizione sono visibili, per un caso di passate parallele nel piano, in figura 5.

La figura mostra schematicamente come la seconda passata, parzialmente sovrapposta alla prima, generi una zona trasformata di ampiezza maggiore rispetto alla precedente. Ciò è dovuto al calore accumulato nel pezzo dalla passata precedente. L'altro effetto di cui bisogna tenere conto è l'effetto di rinvenimento che la seconda passata genera nella struttura martensitica della prima, evidenziato nello schema dalla zona tratteggiata.

I fenomeni di rinvenimento sono ampiamente predicibili mediante modelli empirico sperimentali nel caso di trasformazioni in forno a temperatura costante. Pochissimi sono i lavori che si occupano di predire le trasformazioni delle strutture di tempra in presenza di forti gradienti di temperatura ed i pochi reperibili in letteratura quali [10, 12] richiedono la simulazione della microstruttura e quindi non sono direttamente applicabili a casi industriali.

L'approccio proposto prevede di suddividere il fenomeno in due fasi:

- La ri-austenitizzazione della martensite.
- Il rinvenimento della martensite in presenza di forti gradienti di temperatura.

Si ipotizza che il primo fenomeno sia, in qualche modo, legato a fenomeni di diffusione del carbonio su piccole distanze e quindi esprimibile con la formulazione integrale di equazione 5 simile a quella già proposta per l'austenitizzazione.

$$I_{m \rightarrow a} = \int_{t_{A_{c1}}}^{t_{A_{r1}}} \exp\left(-\frac{Q_{m \rightarrow a}}{RT(t)}\right) dt \quad (5)$$

Il confronto tra il valore calcolato e la soglia $I_{m \rightarrow a, th}$, determinato sperimentalmente per il materiale in esame, permette di determinare la posizione del punto di minima durezza che separa la zona A_1 , soggetta a rinvenimento, dalla zona A_2 soggetta all'austenitizzazione della martensite presente, come visibile in figura 6.

La determinazione dell'entità del rinvenimento all'interno della zona A_2 non può essere ottenuta con le usuali formulazioni empiriche che, a causa degli elevatissimi gradienti

di temperature spaziali e temporali, non possono essere applicate. Per capire quanto il processo si discosti dai normali trattamenti in forno basti pensare che i punti della zona A_1 possono rinvenire, mantenendo la loro identità di strutture di tempra, nonostante picchi di temperature dell'ordine di $1000 \div 1100$ °C. La formulazione proposta per il rinvenimento prevede la definizione di un parametro che tenga conto, istante per istante, degli effetti di rinvenimento causati dalla permanenza nell'intervallo temporale considerato alla temperatura richiesta. L'effetto di rinvenimento passando dal tempo t_k al tempo $t_{k+1} = t_k + \delta t$ può essere calcolato a partire dall'equazione 6 proveniente da [13]. Questa formula fornisce il valore della durezza per un trattamento a temperatura T per un tempo di permanenza t_k .

$$\begin{aligned} H^k &= -74 - 434C - 368Si - 25Mn + 37Ni \\ &- 335Mo - 2235V + \frac{10^3}{P_c(T, t_k)}(260 + 616C + 321Si \\ &- 21Mn - 35Ni - 11Cr + 352Mo + 2354V) \end{aligned} \quad (6)$$

I simboli rappresentano la percentuale di elementi di lega mentre il parametro $P_c(T, t_k)$, come mostrato in equazione 7, tiene conto della coalescenza delle microstrutture risultanti in funzione del ciclo di trattamento termico.

$$P_c(T, t_k) = \left[\frac{1}{T} - \frac{nR}{H_a} \cdot \log \frac{t_k}{t_0} \right]^{-1} \quad (7)$$

H_a è l'entalpia di attivazione della trasformazione microstrutturale durante il rinvenimento, è t_0 rappresenta il tempo unitario di trattamento.

A partire da questa formulazione gli autori propongono l'introduzione del parametro τ_k detto *tempo di rinvenimento* da calcolarsi in quei punti nei quali la martensite presente sale sopra la temperatura di inizio formazione di martensite T_{Ms} . Il parametro τ_k viene calcolato come la diminuzione della durezza che si ottiene supponendo la temperatura costante nell'intervallo di tempo considerato e si ottiene attraverso l'equazione 8.

$$\tau_{k+1} = 1 - \frac{(H^k - H^{k+1})}{H^k} \quad (8)$$

H^k e H^{k+1} sono le durezze che si otterrebbero attraverso l'equazione 6 supposto due trattamenti di temperatura e durata rispettivamente t_k, T_k e t_{k+1}, T_{k+1} .

La durezza attuale della martensite al generico istante t_k viene valutata partendo dalla durezza iniziale H^0 come da equazione 9.

$$H^k = H^0 \prod_{j=0}^k \tau_j \quad (9)$$

5. ESEMPI ED APPLICAZIONI

La validazione dei modelli proposti è stata eseguita attraverso una estesa campagna di prove su leghe ferrose effettuata mediante la sorgente laser da $3kW$ a flusso assiale veloce del laboratorio sperimentale del DIEM.

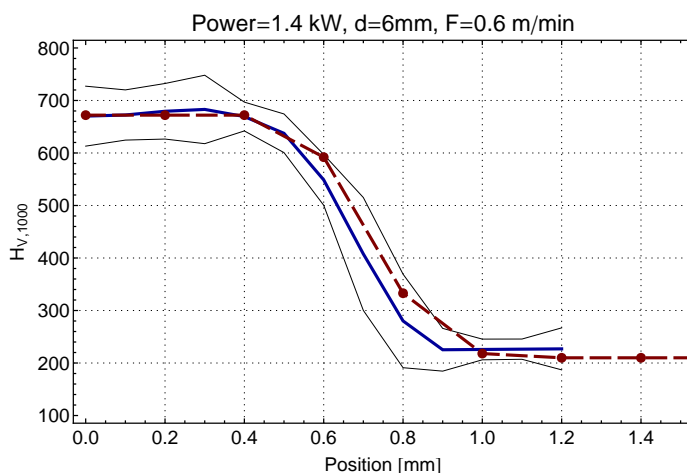


Figure 7: C43: confronto tra durezza sperimentale e simulata nella profondità del cordone.

Acciaio C43

Una delle leghe sulla quale sono stati messi a punto i modelli di austenitizzazione rapida e di rinvenimento è la lega a medio tenore di carbonio C43. Una prima campagna sperimentale è stata eseguita per determinare i valori di $I_{p \rightarrow a, min}$ e $I_{p \rightarrow a, max}$ su questa lega. La campagna ha comportato l'esecuzione di prove di trattamento con potenza variabile su tre livelli tra 1.2 a 1.8 kW e velocità comprese tra 0.3 e 1.1 m/min per un totale di 6 diverse configurazioni focalizzando il fascio laser con un diametro di spot pari a 6.5 mm. Al termine delle simulazioni sono state eseguite delle prove di microdurezza sulla mezzeria del cordone per ottenere il profilo in profondità. Sono state infine condotte una serie di simulazioni variando i due parametri della distribuzione $I_{p \rightarrow a, min}$ e $I_{p \rightarrow a, max}$ e determinando la coppia di valori che minimizza l'errore complessivo tra le microdurezze misurate e quelle simulate sull'insieme dei profili di tutte le configurazioni di prova.

I valori che meglio approssimano i risultati è la coppia $I_{p \rightarrow a, min} = 1.4 \cdot 10^{-4}$ s e $I_{p \rightarrow a, max} = 6.4 \cdot 10^{-4}$ s. Con questi valori sono state condotte delle prove di verifica un cui esempio è mostrato in figura 7 nella quale sono indicati i parametri di prova; la linea tratteggiata rappresenta l'andamento delle durezze simulate e quella continua la durezza misurata con la relativa banda di dispersione a $\pm 3\sigma$.

Sullo stesso tipo di acciaio sono state eseguite delle prove per la determinazione dei valori corretti di $I_{m \rightarrow a}$ in presenza di passate multiple. In particolare, per quanto nelle conoscenze degli autori il software LHS risulta l'unico in grado di analizzare il comportamento del materiale in presenza di una sequenza arbitraria di cicli termici. La figura 8 mostra i risultati misurati trasversalmente al percorso del fascio, ottenuti a 200 μm sotto la superficie per un caso di 3 passate affiancate con una sovrapposizione del 75%. La predizione della posizione della zona di transizione tra la parte rinvenuta e quella austenitizzata risulta decisamente adeguata. Una maggiore incertezza appare nel valutare la diminuzione di durezza nella zona rinvenuta, cosa che richiederà ulteriori studi in futuro. La messa a punto del sistema di simulazione permette di ottimizzare la strategia di processo, ad esempio la figura 9 mostra l'analisi, con relativo confronto sperimentale, di una sequenza di tre passate affiancate realizzate in modo

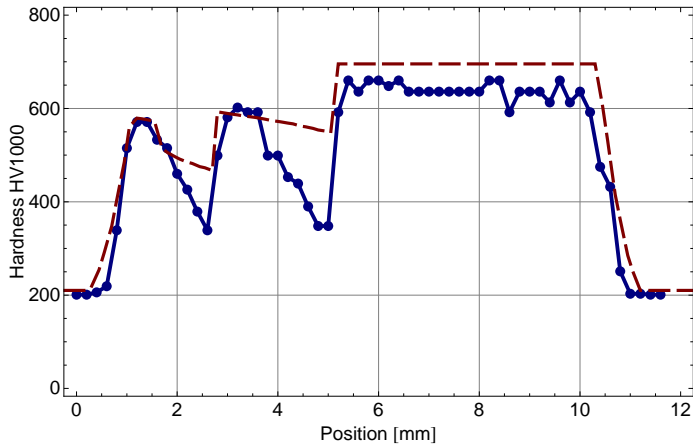


Figure 8: Andamento della durezza in presenza di tre passate sovrapposte.

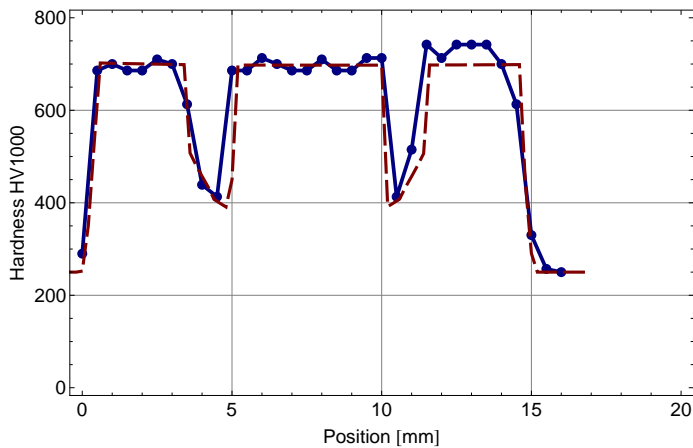


Figure 9: Andamento della durezza in presenza di tre passate sovrapposte alternate.

alternato (prima, terza, seconda...). In pratica nel caso in esame la passata centrale viene realizzata come ultima rispetto alle due laterali si nota un'ottima risposta sia in termini di determinazione della zona riustenitizzata che del valore del rinvenimento.

Acciaio C67

Ulteriori prove sono state condotte su acciaio C67 a tenore di carbonio medio/alto. Anche in questo caso sono state condotte delle prove a singola passata per la determinazione dei parametri ottimali di processo, anche in questo caso sono state condotte una serie di prove a singole passate con 6 differenti configurazioni di parametri di processo. I valori che ottimiz-

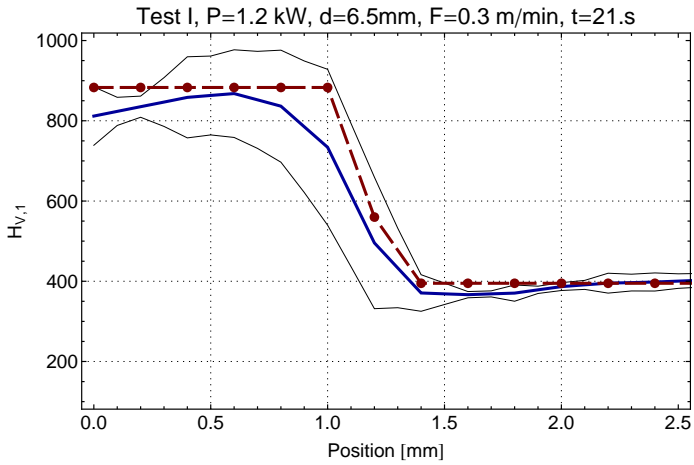


Figure 10: C67: confronto tra durezza sperimentale e simulata nella profondità del cordone a media potenza e bassa velocità.

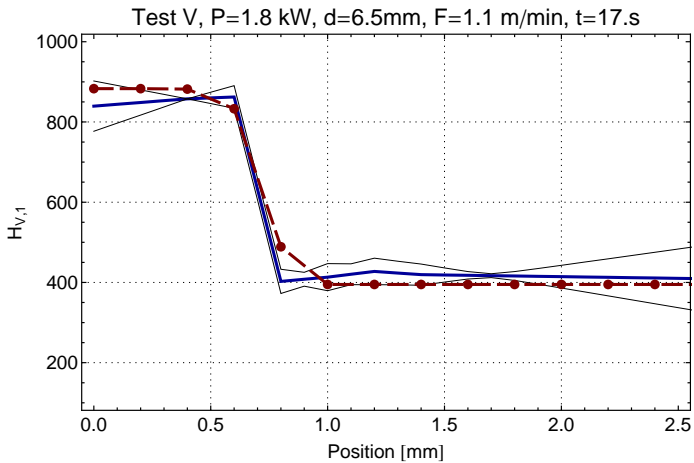


Figure 11: C67: confronto tra durezza sperimentale e simulata nella profondità del cordone con elevata velocità di avanzamento

zani i risultati sono $I_{p \rightarrow a, min} = 8.0 \cdot 10^{-5} s$ e $I_{p \rightarrow a, max} = 2.8 \cdot 10^{-4} s$, i valori sono inferiori rispetto al caso precedente come è ragionevole aspettarsi (la temprabilità del C67 bonificato è maggiore rispetto al C43 ricotto). Le figure 10 e 11 mostrano i risultati del modello al variare della potenza e della velocità di scansione.

BIBLIOGRAFIA

- [1] Tani, G., Orazi, L., and Fortunato, A., 2008. "Prediction of hypoeutectoid steel softening due to tempering phenomena in laser surface hardening". *Annals of Cirp*, **57**(1), pp. 209–212.
- [2] Ashby, M. F., and Easterling, K. E., 1984. "The transformation hardening of steel surfaces by laser beam - i hypo-eutectoid steels". *Acta metallurgica*, **32**(11), pp. 1935–1948.
- [3] Skvarenina, S., and Shin, Y. C., 2006. "Predictive modeling and experimental results for laser hardening of aisi 1536 steel with complex geometric features by a high power diode laser". *Surface & Coatings Technology*, **46**, pp. 3949–3962.
- [4] Jacot, A., and Rappaz, M., 1997. "A two-dimensional diffusion model for the prediction of phase transformation: application to austenization and homogenization of hypoeutectoid fe-c steels". *Acta mater.*, **45**(2), pp. 575–585.
- [5] Reti, T., Fried, Z., and Felde, I., 2001. "Computer simulation of steel quenching process using a multi-phase transformation model". *Computational Materials Science*, **22**, Dec, pp. 261–278(18).
- [6] Denis, S., Farias, D., and Simon, A., 1992. "Mathematical model coupling phase transformations and temperature". *ISIJ International*, **32**(3), pp. 316–325.
- [7] Tani, G., Orazi, L., Fortunato, A., Campana, G., and Cuccolini, G., 2007. "laser hardening process simulation for mechanical parts". In Proceedings of LASE 2007, San Jose - CA, USA.
- [8] Li, W. B., Easterling, K. E., and Ashby, M. F., 1986. "The transformation hardening of steel surfaces by laser - ii hypereutectoid steels". *Acta metallurgica*, **34**(8), pp. 1533–1543.
- [9] Ohmura, E., and Inoue, K., 1989. "Computer simulation on structural changes of hypoeutectoid steel in laser transformation hardening process". *JSME International Journal*, **32**, pp. 45–53.
- [10] Jacot, A., and Rappaz, M., 1998. "Modeling of reaustenization from the perlite structure in steel". *Acta mater.*, **46**(11), pp. 3949–3962.
- [11] Jacot, A., and Rappaz, M., 1999. "A combined model for the description of austenization, homogenization and grain growth in hypoeutectoid fe-c steel during heating". *Acta mater.*, **47**(5), pp. 1645–1651.
- [12] Wang, Y., Denis, B., Appolaire, B., and Archambault, P., 2004. "Modelling of precipitation of carbides during tempering". *Journal of Physics IV*(120), pp. 103–110.
- [13] Committee, A. I. H., 1991. *ASM Handbook, Vol. 04: Heat Treating*, Vol. 4. American Society of Metals, New York.

SIMULAZIONI DI PROCESSI DI COLATA DI COMPONENTI INDUSTRIALI IN ALLUMINIO DI ALTA COMPLESSITA'

Rosario Squatrito

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia

Università di Bologna, Italia

E-mail: rosario.squatrito@mail.ing.unibo.it

Ivan Todaro

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia

Università di Bologna, Italia

E-mail: ivan.todaro@mail.ing.unibo.it

Andrea Morri

SMETEC – Dipartimento di Scienza dei Metalli, Elettrochimica e Tecniche Chimiche

Università di Bologna, Italia

E-mail: a.morri@unibo.it

Luca Tomesani

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Nucleari, Aeronautiche e di Metallurgia

Università di Bologna, Italia

E-mail: luca.tomesani@mail.ing.unibo.it

Abstract. *In questo lavoro è stata portata a termine un'analisi numerica di un processo di colata in conchiglia in gravità di una testa motore V8 in lega A356, attualmente in fase di produzione. Il sistema complessivo di colata, composto da sette parti stampo e un elevato numero di anime in sabbia è stato studiato numericamente, mediante un modello agli elementi finiti costituito da $11.8 \cdot 10^6$ elementi tetraedrici. Le analisi delle fasi di riempimento e solidificazione del getto sono state condotte cercando di replicare in maniera accurata tutte le variabili di processo misurate durante il ciclo di produzione.*

Le caratteristiche microstrutturali, sono state valutate mediante le correlazioni esistenti tra i valori di SDAS e le velocità locali di raffreddamento. Le grandezze calcolate sono state confrontate con i valori misurati da indagini metallografiche su 640 campioni estratti dal getto al termine del processo di colata.

La distribuzione di porosità percentuale è stata quindi stimata numericamente mediante il ricorso al modello numerico del software.

Keywords: *Colata in gravità, A356, Simulazione numerica, SDAS, Porosità da ritiro*

1. INTRODUZIONE

Il mondo dell'automotive vede oggi una concorrenza giocata su standard qualitativi sempre più alti e su caratteristiche strutturali dei componenti sempre più spinte.

Questa richiesta di elevate caratteristiche prestazionali dei componenti, unita alla necessità di miglioramento dei rapporti peso/potenza delle autovetture, porta inevitabilmente alla ricerca di riduzione dei coefficienti di sicurezza nella progettazione, e quindi a tentare di ridurre in maniera determinante le difettologie presenti, facendo ricorso a modifiche sui materiali impiegati, oppure ottimizzando il processo produttivo.

Negli ultimi anni le realtà industriali legate alla produzione di componentistica automotive attraverso processi di fusione, si affidano in maniera sempre più massiccia a strumenti di calcolo in grado di prevedere le dinamiche dei processi di colata, attraverso una simulazione sempre più elaborata, coadiuvata dal costante incremento della potenza computazionale.

L'impiego di questi strumenti di calcolo, attraverso lo studio delle dinamiche di processo, permette di prevedere con buona precisione le caratteristiche finali dei getti colati, correlando eventuali difettosità del pezzo finito a possibili criticità del processo di produzione o ad eventuali errori di progettazione.

E' bene sottolineare come l'insieme dei modelli numerici implementati da un codice di simulazione debba tenere conto della complessità dei fenomeni fisici in gioco, dove rientrano sia gli aspetti fluidodinamici del riempimento della cavità, sia preponderanti fenomeni di scambio termico, oltre agli aspetti metallurgici legati al cambiamento di fase della lega [1].

In quest'ottica l'impiego proficuo di strumenti di calcolo da cui ottenere risultati con un buon livello di affidabilità, viene ad essere influenzato non solo dalla complessità geometrica del componente prodotto ma anche dalla corretta valutazione delle condizioni al contorno che riproducano le reali condizioni di esercizio delle attrezzature.

Risulta quindi evidente come, per le realtà produttive che adottano questi strumenti di calcolo, rivesta una primaria importanza un'accurata fase di "taratura" dello strumento di calcolo da effettuarsi attraverso la misurazione diretta dei parametri di processo.

In questo articolo viene presentato lo studio effettuato su una testa motore V8 da 4.2 litri in lega d'alluminio A356 oggi in produzione presso lo stabilimento Ferrari mediante un processo fusorio di conchiglia in gravità.

L'applicazione della tecnologia di fusione in conchiglia nella produzione di teste motore è sicuramente una delle più complesse nell'industria di fonderia. Infatti, l'alta complessità geometrica del componente richiede l'impiego di numerosi componenti dell'attrezzatura di colata che deve prevedere, oltre alle parti mobili di conchiglia anche numerose anime in sabbia per la realizzazione dei condotti di aspirazione e scarico e i circuiti di raffreddamento.

A tutto questo si aggiunge la richiesta di elevate caratteristiche meccaniche e bassi livelli di porosità, ottenibili attraverso un opportuno controllo della temperatura dei componenti della attrezzatura durante il ciclo produttivo.

Lo scopo di questo lavoro è stato in primo luogo quello di proporre una metodologia di analisi numerica di un processo complesso come quello descritto cercando di emulare in maniera accurata l'insieme di condizioni al contorno del problema.

Infine, attraverso una validazione su dati sperimentali è stato possibile valutare l'affidabilità dei modelli numerici esistenti nella previsione delle caratteristiche microstrutturali del componente, con particolare riferimento ai valori di SDAS, e ai livelli

percentuali di porosità presenti nel getto al termine del processo fusorio.

L'analisi numerica è stata condotta mediante l'impiego del codice commerciale agli elementi finiti Procast V.2008 di ESI Software [2].

L'analisi di riempimento dell'attrezzatura di colata è stata validata mediante osservazione diretta del fenomeno e mediante confronto con dati di processo, con particolare riferimento ai tempi di colata. La successiva fase di analisi della solidificazione è stata quindi portata a termine tramite un opportuna fase di "fine tuning" dei parametri di simulazione, in particolare dei coefficienti di scambio termico per ottenere una convergenza con i valori di temperatura monitorati in alcuni punti dell'attrezzatura durante il processo [3].

Le caratteristiche microstrutturali sono state quindi analizzate sfruttando le correlazioni note tra le velocità locali di raffreddamento e le distanze interdendritiche secondarie (SDAS).

I dati numerici ottenuti dal modello agli elementi finiti sono stati confrontati con i valori ottenuti dalle analisi metallografiche su 640 campioni estratti da un getto di produzione.

I livelli locali di porosità percentuale sono stati calcolati mediante un modello di previsione delle porosità da ritiro e confrontati sui valori misurati nei campioni analizzati.

Sono stati quindi calcolati i coefficienti di correlazione lineare per le grandezze studiate.

2. DESCRIZIONE DEL PROCESSO

Dal punto di vista del fonditore, la realizzazione di un componente complesso come una testa motore, rappresenta senza dubbio una sfida sia per la complessità geometrica del pezzo sia per le caratteristiche funzionali che esso deve svolgere in esercizio.

Infatti il getto estratto al termine del processo fusorio deve essere completamente privo di difetti, e deve presentare lungo determinate superfici, come ad esempio il piatto fiamma, ottime caratteristiche meccaniche per resistere alle alte sollecitazioni strutturali e termiche in esercizio (si tenga presente che le pareti delle camere di combustione raggiungono temperature superiori ai 300 °C). Sommando a queste specifiche anche la necessità di produzione in larga scala, il processo di colata in conchiglia in gravità risulta senza dubbio quello ottimale in quanto caratterizzato dai seguenti benefici:

- Buona tolleranza dimensionale
- Riempimenti con basse velocità che riducono l'inglobamento di aria e ossidi
- Alte velocità di raffreddamento che permettono di ottenere una microstruttura con grani di medio-piccole dimensioni

L'attrezzatura di colata che è stata oggetto di esame è composta da sette parti conchiglia in acciaio, quattro delle quali scorrevoli, da tre supporti in ghisa, e da sette parti anima. Le parti stampo vengono di seguito descritte (Figura 1):

- Il componente inferiore, di seguito definito *fondello*, che descrive la figura del piatto fiamma, risulta attraversato da due distinti circuiti di raffreddamento, il primo longitudinale per limitare la deformazione termica dello stampo, il secondo composto invece da una sequenza di quattro fontane che permettono di controllare localmente la velocità di raffreddamento sul piatto fiamma in corrispondenza dei quattro cilindri.

- Due componenti di fiancata che nella zona interna riproducono la figura destra e sinistra del getto ed esternamente sono solidali a due supporti in ghisa in cui vengono alloggiati dei bruciatori.
- Un carrello frontale e un carrello posteriore che scorrono all'interno di due torri. All'interno di una delle due torri è ricavato il bacino di colata

La sequenza di assemblaggio dell'attrezzatura prima della colata prevede il ramolaggio delle anime in maniera manuale dall'operatore e il loro inserimento tramite braccio automatizzato all'interno della conchiglia nella fase di apertura.

Il sistema di colata prevede una doppia alimentazione, rispettivamente a sinistra e a destra dell'asse longitudinale del getto. Due bastoni di colata collegano il bacino superiore a due canali laterali che vanno ad alimentare gli attacchi di colata. Le sezioni e le lunghezze dei due canali di alimentazione degli attacchi risultano di dimensioni diverse, così come vi è una asimmetria del numero di attacchi di colata rispetto all'asse longitudinale. Nella zona di accesso ai canali di colata vengono alloggiati due filtri ceramici. La lega impiegata viene ricavata dalla A356, raffinata mediante l'aggiunta di Ti e B e modificata mediante Sr.

La temperatura del metallo all'atto della colata viene tenuta sotto controllo monitorando la temperatura del bagno nel forno di iniezione. Attraverso un tubo pescante verticale posto nel forno, alimentato da una rampa di pressione di iniezione opportunamente tarata, si effettua un riempimento del bacino di colata con un bilanciamento della massa di metallo versata nel bacino e quella che fluisce nella conchiglia. In questo modo si cerca di mantenere un livello costante di altezza di pelo libero all'interno del bacino.

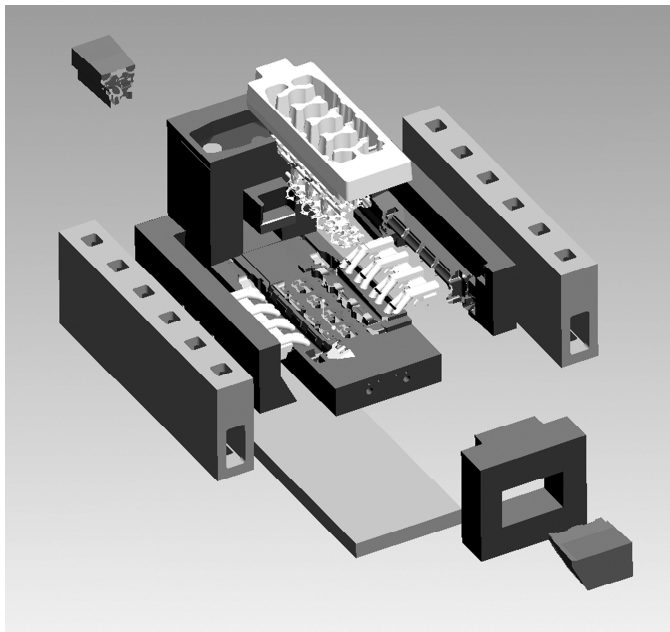


Figura 1. Attrezzatura di colata

Durante la fase di solidificazione vengono controllati i valori di temperatura in alcuni punti mediante termocoppie collegate al pannello di controllo della stazione di colata. I valori di temperatura misurati dalle sonde servono alla termoregolazione della conchiglia e comandano in maniera semiautomatica l'apertura dei canali di raffreddamento del fondello e l'alimentazione dei bruciatori posti lungo le fiancate.

Al termine della solidificazione le parti mobili dell'attrezzatura vengono aperte e il getto viene estratto per essere quindi marcato e quindi avviato alle successive fasi di lavorazione.

3. ANALISI NUMERICA

L'approccio numerico al problema ha visto differenti fasi di studio durante le quali è stato effettuato un confronto tra i risultati numerici e l'osservazione diretta dei dati in stabilimento, in particolare:

1. La prima fase di analisi ha permesso di determinare le condizioni di esercizio dell'attrezzatura, calcolando gli andamenti della distribuzione di temperatura nelle parti stampo durante le fasi di ciclaggio termico. In questa fase la distribuzione di temperatura al termine del singolo ciclo è stata utilizzata come condizione iniziale per il calcolo dei valori di temperatura nel ciclo successivo
2. Imponendo come valori di temperatura iniziali delle varie parti stampo quelli ottenuti al termine della fase 1 è stata quindi effettuata un'analisi della fluidodinamica di riempimento della cavità
3. Infine, imponendo come condizioni iniziali di temperatura sia del getto che dell'attrezzatura, quelle ottenute al termine della fase 2, è stata quindi portata a termine l'analisi del raffreddamento e della solidificazione del getto.

Poiché sia l'evoluzione della microstruttura, sia la formazione delle porosità legate ai fenomeni di ritiro di solidificazione sono entrambi aspetti fortemente correlati alle condizioni di raffreddamento del getto è stato necessario includere nel dominio di calcolo tutti i componenti dell'attrezzatura per tenere conto della totalità dei fenomeni di scambio termico.

Il modelli geometrici sia del getto che dei vari componenti dell'attrezzatura precedentemente descritti sono stati quindi discretizzati secondo elementi tetraedrici con dimensione caratteristica più piccola nelle zone del getto e di stampo in prossimità dell'interfaccia metallo/acciaio. La dimensione degli elementi viene riportata in tab. 1.

Tabella 1. Dimensioni caratteristiche della griglia di calcolo

Componente	Dimensioni caratteristiche elementi min-max	Numero totale di elementi
Getto	1mm -10 mm	2.9×10^6
Parti stampo	1 mm – 10 mm	6.0×10^6
Anime	3 mm – 5 mm	2.3×10^6
Supporti bruciatori	10 mm – 20 mm	5.2×10^5
TOTALE		$11,8 \times 10^6$

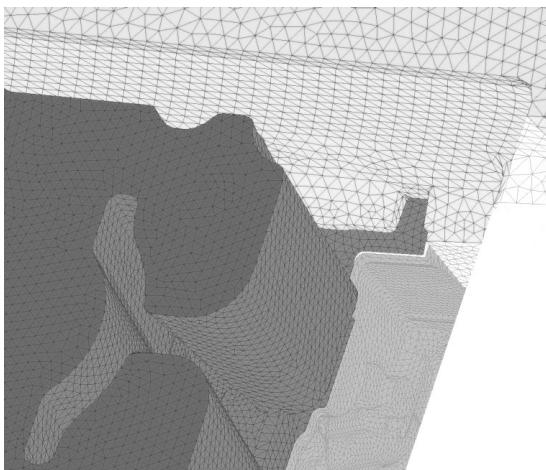


Figura 2. Griglia di calcolo lungo l'interfaccia metallo-stampo

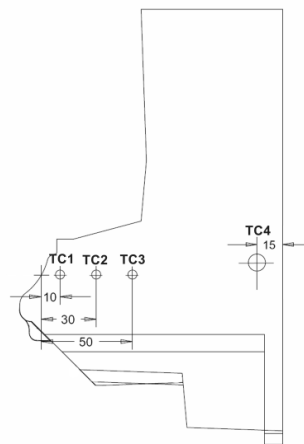


Figura 3. Posizionamento sonde di misura

La necessità di valutare in maniera accurata i fenomeni di scambio all'interfaccia ha richiesto il posizionamento dei nodi di elementi lungo le superfici di contatto, sia tra metallo e stampo che tra le varie superfici di contatto tra le diverse parti dell'attrezzatura (Fig. 2).

I nodi lungo le superfici di interfaccia sono stati quindi considerati coincidenti, e appartenenti a due o più domini di calcolo. Questa scelta è stata dettata dalla necessità di evitare operazioni di interpolazione nel calcolo dei flussi attraverso gli elementi di interfaccia tra fuso e stampo, portando però ad un inevitabile incremento del numero di elementi anche nelle zone del dominio non interessate dalla fluidodinamica di riempimento.

Questa condizione di coincidenza dei nodi degli elementi all'interfaccia, e la necessaria presenza di una transizione graduale verso una griglia più rada nelle zone esterne dell'attrezzatura, ha inevitabilmente influito sul numero totale di elementi del dominio di calcolo e sull'onere computazionale.

Per valutare la corretta determinazione dei coefficienti termici di interfaccia (HTC) tra metallo liquido e stampo è stata condotta una campagna di acquisizione della temperatura in sei punti delle parti conchiglia, con il fine di valutare i dati numerici calcolati durante la fase di ciclaggio termico dell'attrezzatura.

I punti di misura rilevati attraverso termocoppie sono raffigurati in fig. 3 e vengono qui di seguito elencati:

- TC1, TC2, TC3 sul componente di fiancata destro, a distanza rispettivamente di 10 mm, 30 mm, e 50 mm dall'interfaccia di contatto con il metallo. Ciò ha permesso di valutare i gradienti termici nella zona interna della fiancata e di validare i risultati numerici ottenuti, dopo una attività di messa a punto dei coefficienti di scambio termico impiegati nell'analisi numerica.
- TC4, TC5 rispettivamente nella parte destra (aspirazione) e sinistra (scarico), rispettivamente a 15 mm e a 10 mm dall'interfaccia. La misura di questi valori ha permesso di determinare i coefficienti di scambio in corrispondenza

delle interfacce esterne delle fiancate che sono direttamente influenzate dalla presenza dei bruciatori

- TC6 posizionata nel fondello, vicino il circuito di raffreddamento. Questa misura ha permesso di stimare l'effetto dei canali di termoregolazione e di determinare i coefficienti di convezione da applicare alla superficie dei canali.

I valori di temperatura sono stati acquisiti mediante un sistema multiplexer a 16 canali con una frequenza di acquisizione di 0.5 Hz.

4. RISULTATI

Nella prima fase di studio, i risultati ottenuti nella prima fase di analisi della regimazione termica dell'attrezzatura sono stati confrontati con le curve dei valori di temperatura misurati in dipendenza del tempo.

Il confronto degli andamenti delle curve mostra una corretta stima numerica della distribuzione di temperatura nelle componenti stampo che sono stati monitorati. In tre nodi della griglia di calcolo, corrispondenti alle termocoppie TC1, TC2, TC3, si osserva che le escursioni termiche della fiancata a regime sono previste con una buona corrispondenza dei valori di minimo e di massimo della temperatura in tutti i punti misurati.

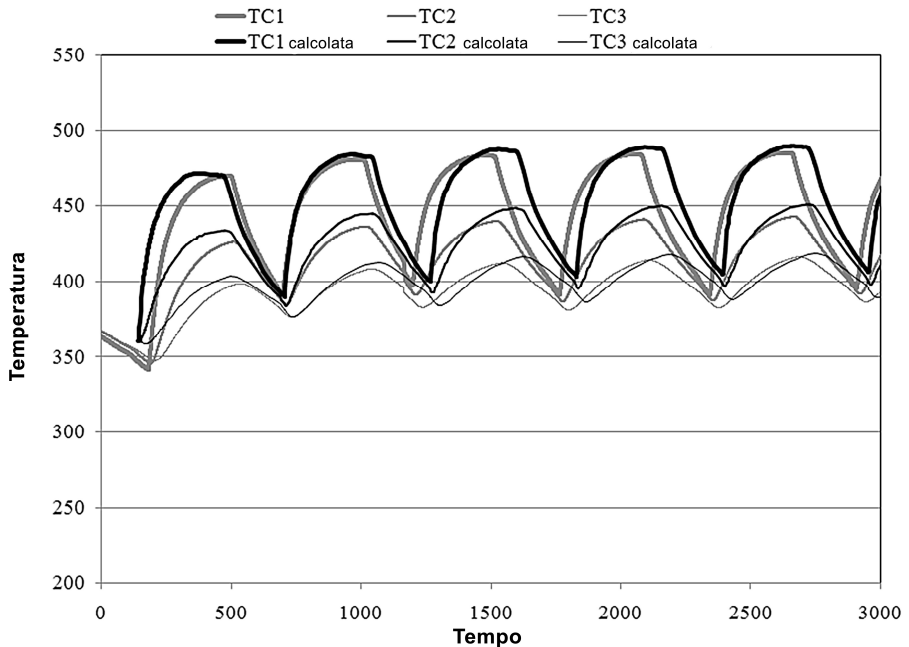


Figura 4. Andamento dei valori di temperatura misurati e calcolati nei punti di fiancata

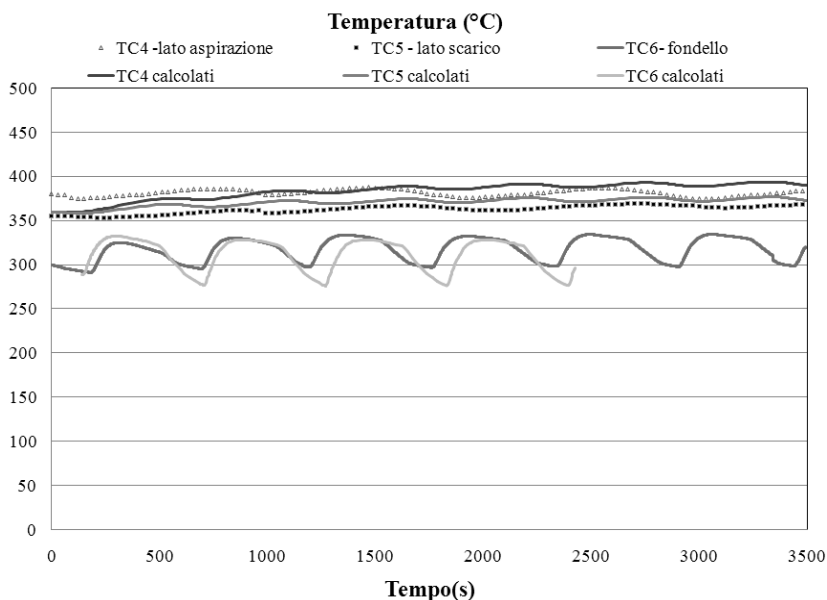


Figura 5. Temperatura di esercizio - zone esterne fiancate e fondello

Inoltre, le pendenze delle curve misurate e calcolate mostrano una buona corrispondenza sia durante la fase di chiusura della conchiglia (riscaldamento), sia nella fase di apertura (raffreddamento).

Nella fig. 5 viene riportato il confronto tra i valori di temperatura previsti e misurati nei punti dove sono state collocate le termocoppie TC4, TC5, in prossimità dei bruciatori, e TC6 nel fondello. Osservando l'andamento di questi dati di temperatura si è limitata l'analisi alla semplice determinazione degli intervalli di esercizio delle zone più esterne dell'attrezzatura che sono difficilmente riproducibili in simulazione in quanto fortemente influenzate dall'azione manuale dell'operatore e sono quindi caratterizzate da scarsa ripetibilità tra un ciclo e l'altro.

Nella seconda fase di studio, è stata quindi analizzata la corrispondenza tra i risultati numerici ottenuti e quelli basati sulle indagini metallografiche effettuate su un getto estratto al termine di un ciclo di produzione, concentrando l'attenzione sul valore di SDAS rilevato e sulla distribuzione percentuale media di porosità. Sono state quindi estratte quattro sezioni dal getto esaminato e ogni sezione è stata divisa in macrosezioni. Le macrosezioni sono state quindi ripartite in celle campione tramite una griglia regolare ottenendo circa 150-180 campioni per sezione.

Mediante la scrittura di un apposito algoritmo di elaborazione dei dati numerici ottenuti sull'intero dominio del getto è stato possibile estrarre il valore medio delle variabili calcolate in tutti i nodi di griglia del modello numerico appartenenti al volume della singola cella campione estratta dal getto.

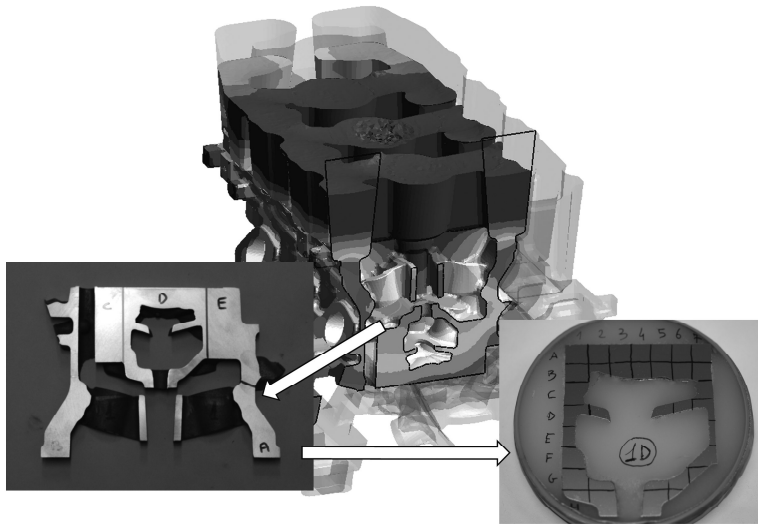


Figura 6. Zone di estrazione dei campioni

In particolare è stata individuata una corrispondenza biunivoca tra valore medio di SDAS e livelli di porosità percentuale calcolati sulla griglia di calcolo FEM e il valore medio delle stesse grandezze osservate sulla griglia di campionatura nelle indagini metallografiche .

Vengono qui di seguito presentati i risultati ottenuti su due delle quattro sezioni analizzate (sezione 1 e sezione 4). I valori di temperatura nel tempo ottenuti dall'analisi del transitorio di solidificazione del getto in un ciclo di produzione considerato in condizioni medie di esercizio sono stati elaborati per ottenere il valore di SDAS, facendo ricorso alla formulazione seguente [4]:

$$SDAS = \left[\left(166 \cdot \frac{\Gamma D \ln(c_{eut} / c_0)}{m(1-k)(c_0 - c_{eut})} \right) \cdot t_f \right]^{1/3} \quad (1)$$

Dove t_f è il tempo medio di solidificazione, calcolato tra il valore di temperatura di liquidus e quello corrispondente alla prima trasformazione eutettica, D è il coefficiente di diffusione del soluto nel liquido, Γ è il coefficiente di Gibbs-Thomson, m è la pendenza della curva di liquidus.

In fig. 7 vengono riportate le matrici dei valori medi di SDAS calcolati e misurati nella sezione 1 in ciascuna cella campione.

Osservando la distribuzione dei valori nelle diverse zone di sezione si osserva una corrispondenza globalmente accettabile, con una corretta previsione degli andamenti lungo ciascuna riga della matrice. Questa corrispondenza suggerisce una buona accuratezza raggiunta nella determinazione delle corrette condizioni al contorno del problema studiato.

	E	D	C		B	A	1	2	3	4	5	6	7	A	B	C	D		SDAS
1			33		40	41	40	40	38	38	38	38	38	39	35	29		7	29
2	36	37	37		44	44	43	43	42	42	42	42	42	41	39	38		6	31
3			41		47	47	45						44	43	42	41		5	33
4		43	43		49	49	48	49			48	47	46	46	44	41	37	4	36
5			36		50	51	50						48	49	48	44	46	3	38
6			35				51					50				48	47	2	41
7			35	36			53	55			55	53	51			48	44	1	48
6			37	39	39		54	55			56	55			47	46	45	6	45
5			37	39	39										44	41	39	5	48
4			32	35	37											38	36	4	50
3	30	32	34								57				37	36	32	3	53
2	31	34	34								57					37	34	2	55
1	34	36	36								55					40	37	1	57
	A	B	C	D	E					F				D	C		B	A	

	E	D	C		B	A	1	2	3	4	5	6	7	A	B	C	D		SDAS
1			26		34	36	37	34	35	36	36	36	35	31	27	26		7	26
2	30	32	34		42	46	42	39	42	44	43	44	40	43	38	33		6	28
3			38		52	49	49						50	48	43	36		5	31
4		37	36		58	47	51	44			47	53	50	44	48	45	45	4	33
5			36		56	48	49						53	50	43	48	46	3	36
6			35				52					49			49	45	2	38	
7			36	39			51	54			63	53	52		47	45	1	40	
6			37	40	39		55	59			56	55			45	37		6	43
5			40	39	38										38	37		5	45
4		41	42	40											38	39		4	48
3	33	36	39								31				32	36	29	3	50
2	34	38	38								32					36	34	2	52
1	39	44	40								29					40	40	1	55
	A	B	C	D	E					F				D	C		B	A	

Figura 7. a) Valori SDAS calcolati - sezione 1; b) SDAS misurati - sezione 1

La fig.8 mostra inoltre la correlazione riscontrata per tutte le celle della sezione 1 dove in ascissa vengono riportati i valori misurati e in ordinata quelli calcolati. Il coefficiente globale di correlazione calcolato dalla distribuzione dei dati di sezione 1 è pari a $R=0.605$.

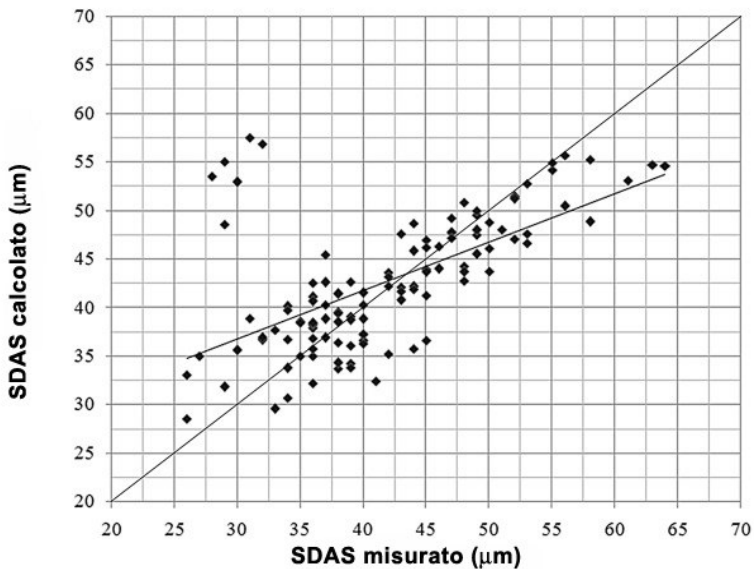


Figura 8. Correlazione valori di SDAS misurati/calcolati - sezione 1

Come si può osservare confrontando l'andamento della linea di correlazione dei dati riportati con la bisettrice del grafico, si riscontra una generale tendenza a sovrastimare i valori SDAS nelle zone a micro struttura più fine e, al contrario, a sottostimare le zone ad alto valore di SDAS (>40 micron).

Questa tendenza è inoltre confermata dalla presenza nel grafico di fig.8 di alcuni punti fortemente discostati dalla bisettrice e che sono corrispondenti a due zone della sezione 1 aventi spessori molto sottili ed interamente circondate da un'anima centrale in sabbia, osservabili nella zona inferiore delle matrici di fig. 7.

Una giustificazione dello scostamento riscontrato in questi precisi punti della sezione può ricercarsi nel loro basso modulo termico e quindi nei bassi tempi di solidificazione ad essi associati. A tal riguardo, si ritiene che l'andamento costante assunto per i valori di HTC di interfaccia possa indurre una percentuale rilevante di errore nelle zone di sezione particolarmente sottile, dove risulta necessario tenere conto della variabilità del coefficiente di scambio termico in funzione del tempo.

Se si escludono dal calcolo del coefficiente globale di correlazione i punti che hanno mostrato maggiori scostamenti, si ottiene un valore $R=0.86$, che, secondo gli Autori, è maggiormente rappresentativo della correlazione globale riscontrata su tutta la sezione e che risulta allineato con i valori di correlazione riscontrati lungo le altre sezioni investigate.

Nelle fig.9-10 si riportano a titolo di esempio i grafici di distribuzione dei valori calcolati e misurati nella sezione 4, dove è stato ottenuto un indice $R=0.96$.

Per valutare la percentuale di porosità nelle sezioni indagate è stato impiegato il modello numerico implementato nel codice commerciale per prevedere i fenomeni di ritiro del getto durante la solidificazione.

La stessa procedura impiegata per il confronto sui dati di SDAS è stata quindi applicata all'analisi delle distribuzioni percentuali di porosità, ricercando le correlazioni tra i dati misurati e calcolati.

	A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T	U	V	SDAS
1				30	34	35	35	35				36	34	35	34		24				22
2		35	35	35		37	39	39	41	42	41	39	39	38	37	35	31				25
3				37				42	44	45	45	43		43	40	38	37				28
4			43	42	43	45	45	47	49	49	49	47	45	44	44	43	42				31
5						50	49		51	51	51		49	47	46	45	45	45	45	45	34
6					51	51			53	53	53		52	52	51			45	44	42	36
7				44	48	53	54	55	55	55	55	54	54	53	53	53					42
8				43				55	56	57	56		55	55	53	52	52				39
9				43	45	47	50		57	57	57		55	55	53	52	50				42
10				43	45	47	52	57	58	58	58	59	57	54	53	51	50				45
11				42	46	48	51	56	59	59	59	59	58	55	54	52					48
12				37	42	46	50	55	58	60	60	61	60	59	57	54	52	48			51
13	22	29	37	43	48	52	56	59	61	61	61	61	60	59	56	53	49	44			54
14	26	31	39	46	51	54	58	61	63	63	63	62	62	61	59	56	52	47			57
																					60

	A	B	C	D	E	F	G	H	I	L	M	N	O	P	Q	R	S	T	U	V	SDAS
1				26	26	32	32	31				30	32	32	28		28				23
2		30	30	29		35	36	35	39	40	36	38	40	37	35	32	28				26
3				38				40	44	44	47	41		49	42	41	34				30
4			43	38	42	46	46	43	50	47	50	49	49	45	43	41	38				33
5						47	49		55	55	54		49	48	44	42	41	42	44		37
6					56	51			56	58	57		56	57	49			43	39	38	40
7				45	51	53	54	58	59	56	64	59	53	58	53	51					44
8				40				54	61	60	58		56	55	52	53	50				47
9				46	51	53	54		57	61	58		58	53	52	57	45				50
10				47	48	51	56	62	59	63	63	62	67	60	53	49	42				54
11				41	50	51	61	60	63	61	63	60	66	64	56	51					57
12				45	49	55	52	62	66	67	61	67	65	66	63	62	55	49			61
13	23	26	41	49	50	61	68	64	62	69	71	66	68	68	66	57	53	48			64
14	26	27	43	56	58	67	65	66	69	67	65	68	69	69	67	60	59	51			68

Figura 9. Valori di SDAS - sezione 4 a) calcolati b) misurati

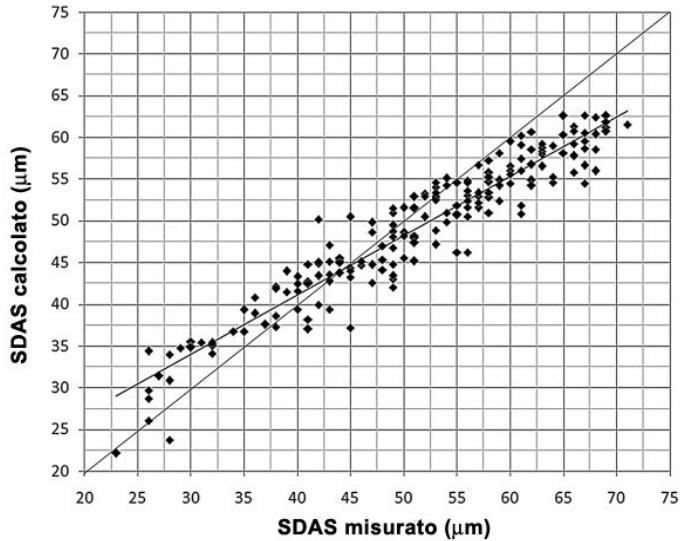


Figura 10. Correlazione valori di SDAS misurati/calcolati - sezione 4

Osservando i valori riportati nelle matrici di distribuzione della porosità percentuale (fig. 11) e le fig. 12-13 riportanti i valori di porosità percentuale misurati e calcolati nella sezione 1 e nella sezione 4 si evidenziano forti dispersioni dei punti e bassi coefficienti di correlazione ($R=0.58$ per la sezione 1 e $R=0.04$ per la sezione 4).

	E	D	C	B	A	1	2	3	4	5	6	7	A	B	C	D	7	Porosity %
1			0.07	0.08	0.08	0.10	0.12	0.12	0.11	0.10	0.10	0.08	0.05	0.05	0.03			0.00
2	0.14	0.07	0.06	0.04	0.04	0.03	0.09	0.39	0.48	0.28	0.12	0.04	0.04	0.05	0.06		6	0.05
3			0.15	0.03	0.02	0.02				0.02	0.03	0.04	0.06	0.06			5	0.10
4		0.34	0.33	0.02	0.01	0.18	0.53			0.39	0.64	0.07	0.04	0.05	0.05	0.01	4	0.15
5			0.11	0.30	0.03	0.00				0.02	0.06	0.09	0.07	0.17		3	0.20	
6			0.23	0.19	0.24	0.03				0.04			0.52	0.24		2	0.25	
7			0.19	0.17	0.24	0.04	0.09			0.18	0.07	0.07	0.34	0.16	1	1	0.30	
6			0.17	0.13	0.16	0.36				0.15	0.06		0.43	0.23	0.20		6	0.35
5			0.13	0.13	0.31					0.15	0.06		0.15	0.10	0.12		5	0.40
4		0.11	0.09	0.09									0.11	0.11			4	0.45
3	0.10	0.13	0.31			0.60				0.64			0.32	0.25	0.11	3	0.49	
2	0.10	0.12	0.32			0.57				0.60			0.16	0.09	2	2	0.54	
1	0.06	0.05	0.05			0.33				0.32			0.04	0.03	1	1	0.59	
	A	B	C	D	E	F				E			D	C	B	A		

	E	D	C	B	A	1	2	3	4	5	6	7	A	B	C	D	7	Porosity %
1			0.25	0.24	0.19	0.24	0.31	0.34	0.27	0.29	0.26	0.19	0.19	0.21	0			0.00
2	0.34	0.34	0.33	0.31	0.26	0.52	0.66	0.53	0.49	0.51	0.38	0.27	0.18	0.21	0.24		6	0.10
3			0.48	0.32	0.18	0.37				0.55	0.18	0	0.24				5	0.21
4		0.41	0.81	0.61	0.19	0.31	0.92			0.73	0.99	0.37	0.14	0.18	0.16	0.12	4	0.31
5			1.12	0.85	0.38	0.28				0.26	0	0.18	0.28	0.24	3	3	0.42	
6			0.66	0.39		0.39				0.62	0.54		0.62	0.61	2	2	0.52	
7			0.38	0.37		0.36	0.38			1	0.35	0.59	0.59	0.88	1	1	0.62	
6			0.54	0.54	0.33	0.66	1.26	0.55		0.63	0.73		0.67	0.31	0.66		6	0.73
5			0.47	0.31	0.66					0.85	0.32	0.46	0.85	0.32	0.46		5	0.83
4		0.32	0.19	0.33						0.36	0.76		0.62	0.39	0.23	3	1.04	
3	0.46	0	0.79			0.31				1.09			0.62	0.39	0.23	3	1.04	
2	0.39	0.33	0.91			1				1.35			0.52	0.25	2	2	1.14	
1	0.32	0.32	0.54			0.52				1.01			0.81	0.26	1	1	1.25	
	A	B	C	D	E	F				E			D	C	B	A		

Figura 11. Porosità percentuale - sezione 1 a) misurata b) calcolata

Porosità da ritiro - Sezione 1

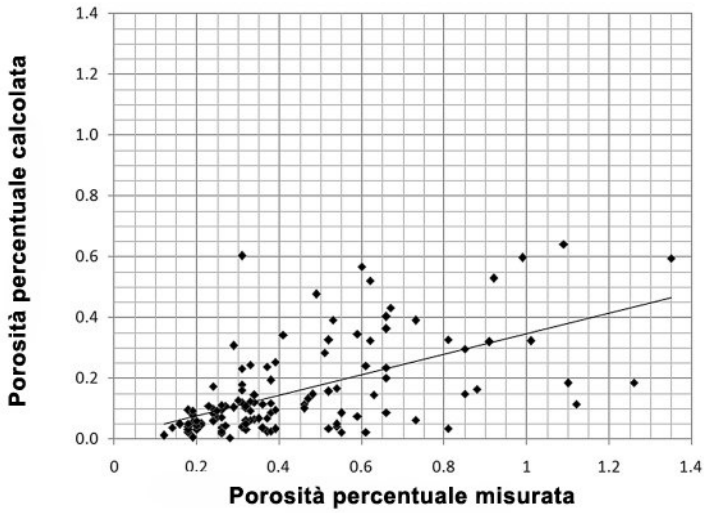


Figura 12. Correlazione dei livelli di porosità percentuale – sezione 1

Porosità da ritiro - Sezione 4

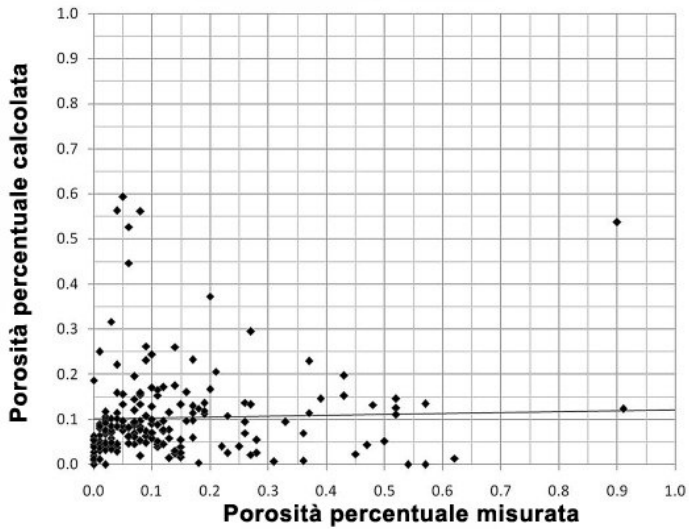


Figura 13. Correlazione dei livelli di porosità percentuale - sezione 4

E' bene a tal proposito evidenziare come i dati ottenuti nel calcolo e nelle analisi di laboratorio mostrano valori massimi di porosità percentuale di poco superiore all'1% sensibilmente inferiori ai valori considerati critici (10%- 15%) per la definizione di scarti nei comuni processi di fonderia.

5. CONSIDERAZIONI CONCLUSIVE

In questo lavoro è stata portata a termine un'analisi numerica di un processo di colata in gravità di una testa motore V8 in alluminio A356, attualmente in produzione. Sia le fasi di riempimento che quelle di solidificazione sono state studiate cercando di replicare in maniera accurata le esatte condizioni di esercizio dell'attrezzatura che sono state misurate durante un normale ciclo di produzione.

Le caratteristiche microstrutturali, in particolare i valori locali di SDAS misurato su un numero complessivo di 640 campioni, sono stati confrontati con i valori calcolati con un metodo agli elementi finiti impiegando le correlazioni esistenti tra le grandezze analizzate e le velocità locali di raffreddamento. I fattori di correlazione ottenuti tra i valori misurati e calcolati sono risultati compresi nell'intervallo di 0.85-0.96.

Il contenuto di porosità percentuale è stato quindi valutato ricorrendo al modello di previsione della porosità da ritiro implementato nel codice commerciale.

I fattori di correlazione trovati nella previsione dei livelli di porosità sono risultati compresi tra 0.04 e 0.58. Si sottolinea come il livello massimo di porosità percentuale misurato risulta sempre inferiore al 2%, notevolmente inferiore ai limiti di sensibilità segnalati dal produttore del software.

Ringraziamenti

Si ringrazia Ferrari Spa , in particolare l'Ing. Gianluca Pivetti per la collaborazione dimostrata, l'Ing. Gianmaria Fulgenzi e l'Ing Mario Camozzi per gli ausili tecnici messi a disposizione.

BIBLIOGRAFIA

- [1] John Campbell, 2003, *Casting*, Oxford, UK, Elsevier Science Ltd
- [2] ProCAST User Manual V. 2008, 2008, 479-480
- [3] Meneghini A and Tomesani L., 2005, "Chill material and size effects on HTC evolution in sand casting of aluminum alloy", *Journal of Materials Processing Technology*, 162-163, pp. 534-539
- [4] W. Kurz and D.J.Fisher, 1998, *Fundamentals of solidification*, Switzerland, Trans Tech Publications Ltd
- [5] Pequet Ch., Gremaud M. and Rappaz M., 2002, "Modeling of Microporosity, Macroporosity, and Pipe-Shrinkage Formation during the Solidification of Alloys Using a Mushy_zone Refinement Method: Applications Aluminium Alloys", *Metallurgical and material transactions*, 33A, pp. 2095-2106.

REVIEW AND COMPARISON OF THE METRICS PROPOSED FOR MEASURING THE DISTANCE BETWEEN TWO RIGID-BODY POSES

Raffaele Di Gregorio
Department of Engineering
University of Ferrara, Italy
E-mail: raffaele.digregorio@unife.it

Abstract. *The concept of distance between two rigid-body poses is central for lots of applications: path planning, positioning precision, mechanism synthesis, etc.. Many formulas have been devised to match the needs of real motion tasks, and some milestones have been put on the road that drives toward the correct definition of distance metrics useful for design and application purposes. In this context, this paper presents a critical review of the literature on the metrics proposed for measuring the distance between two rigid-body poses.*

Keywords: *SE(3), metrics, positioning precision, mechanism synthesis*

1. INTRODUCTION

The concept of distance between two rigid-body poses (positions and orientations) is central for lots of applications: path planning, positioning precision, mechanism synthesis, etc.. It is implemented by introducing a real-valued mapping, $\rho(\mathbf{x}_1, \mathbf{x}_2)$, of two six-tuples (i.e., $\rho: \mathbb{R}^6 \times \mathbb{R}^6 \rightarrow \mathbb{R}$), \mathbf{x}_1 and \mathbf{x}_2 , (\mathbf{x}_1 and \mathbf{x}_2 identify the two rigid-body poses) which is positive definite, symmetric and satisfies the triangle inequality. Such a mapping is named distance metric.

Defining a distance metric in the configuration space (c-space¹) of a rigid-body is a difficult problem many researchers have dealt with. It has been mainly addressed according to the two following approaches: (i) the attempt to introduce a distance metric directly in the c-space [1-8]; (ii) the attempt to approximate the displacement in the c-space with a spherical or hyper-spherical displacement, and, then, to use a distance metric of the spherical, SO(3), or hyper-spherical, SO(N), space [9-16].

These investigations highlighted that: (1) no bi-invariant² Riemannian metric can be defined in the special Euclidean group, SE(3) (see [4] for references); (2) the size of the rigid body must be considered for defining meaningful distance metrics [2, 3]; (3) bi-invariance is not necessary to define meaningful distance metrics [6].

¹ The coordinates of a point of the c-space are the parameters that locate the rigid-body pose represented by that c-space point.

² A metric is said bi-invariant if it depends neither on the choice of the reference system fixed to the rigid body (body frame), nor on the choice of the reference system fixed to the observer (inertial frame). Moreover, a metric is said left-invariant (right-invariant), if it does not depend on the choice of the inertial frame (the body frame).

In this paper, a critical review of the most known distance metrics proposed in the literature is presented together with a technique, recently proposed by the author [17], for generating distance metrics.

The paper is organized as follows: section 2 will analyze the direct introduction of distance metrics in SE(3) (approach (i)), section 3 will discuss the indirect introduction of distance metrics in SE(3) (approach (ii)), and the last sections will present the author proposal together with the conclusions.

2. DIRECT INTRODUCTION OF DISTANCE METRICS IN SE(3)

The natural way to define a distance metric (see [4]) in a differentiable manifold, \mathbf{M} , consists in implementing the following three steps: (a) definition of a Riemannian metric³, (b) determination of the geodesic⁴ induced by the Riemannian metric, defined in the previous step, and (c) calculation of the explicit expression of the length⁴ of the geodesic that joins two generic points, say \mathbf{X}_1 and \mathbf{X}_2 , of \mathbf{M} . In this procedure, the distance metric is the explicit expression of the geodesic length, determined in the last step.

The special Euclidean group, SE(3), is the set that collects all the rigid-body displacements. Its generic element, \mathbf{X} , can be represented by an homogeneous transformation, that is $\mathbf{X} \triangleq \begin{bmatrix} \mathbf{R} & \mathbf{b} \\ \mathbf{0}^T & 1 \end{bmatrix}$ where (see Fig. 1) \mathbf{R} is the rotation matrix that identifies the rotation of the body frame from the coincidence with the inertial frame to the actual orientation, \mathbf{b} is the position vector of the body-frame origin, B, measured in the inertial frame, and $\mathbf{0}$ is the null vector.

With this representation of \mathbf{X} , the law of composition that gives SE(3) the structure of a group is the product of matrices. Moreover, since a rigid-body path, $\mathbf{X}(t)$, can be easily defined together with its derivative⁵, $\dot{\mathbf{X}}(t)$, SE(3) has the structure of a differentiable manifold (i.e., it is a Lie group).

Hereafter, the twist⁵ will be used to identify $\dot{\mathbf{X}}$, and to write the explicit expressions involving $\dot{\mathbf{X}}$ (i.e., if it is not differently specified, it will be assumed $\dot{\mathbf{X}} \equiv \{\boldsymbol{\omega}^T, \dot{\mathbf{b}}^T\}^T$ in the inertial frame and $\dot{\mathbf{X}} \equiv \{\boldsymbol{\omega}^T, (\dot{\mathbf{b}} - \tilde{\boldsymbol{\omega}}\mathbf{b})^T\}^T$ in the body frame). With this choice, a bi-invariant Riemannian metric could be assigned through the following relationship

³ A Riemannian metric is a smooth assignment of an inner product, $\langle \cdot, \cdot \rangle$, to the tangent space at each point of the manifold.

⁴ The geodesic is the curve, $\mathbf{X}(t) \subset \mathbf{M}$ with $a \leq t \leq b$, that minimizes the integral $L = \int_a^b \sqrt{\langle \dot{\mathbf{X}}, \dot{\mathbf{X}} \rangle} dt$. L is the length of the curve $\mathbf{X}(t)$ measured by using the inner product $\langle \cdot, \cdot \rangle$.

⁵ In SE(3), the tangent space at \mathbf{X} is the set of matrices $\dot{\mathbf{X}} \triangleq \begin{bmatrix} \tilde{\boldsymbol{\omega}}\mathbf{R} & \dot{\mathbf{b}} \\ \mathbf{0}^T & 0 \end{bmatrix}$ where $\tilde{\boldsymbol{\omega}}$ is the skew-symmetric matrix associated to the rigid-body angular velocity, $\boldsymbol{\omega}$. Such a tangent space together with the sum of matrices is a vector space which is isomorphic both to the tangent space at the identity element of SE(3) (called Lie algebra and denoted se(3)), and to the set of rigid-body twists, that is, $\{\boldsymbol{\omega}^T, \dot{\mathbf{b}}^T\}^T$, when referred to the inertial frame (body-fixed velocity), or $\{\boldsymbol{\omega}^T, (\dot{\mathbf{b}} - \tilde{\boldsymbol{\omega}}\mathbf{b})^T\}^T$, when referred to the body frame (space velocity). Thus, the derivative at \mathbf{X} can be uniquely determined either by an element of se(3) or by the corresponding twist.

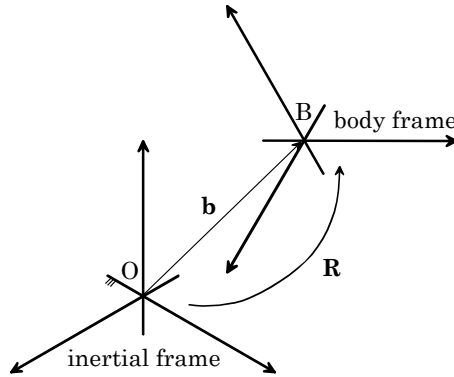


Figure 1. Notations: the inertial frame is fixed to the observer; the body frame is fixed to the rigid body that moves; \mathbf{R} is the rotation matrix that identifies the rotation of the body frame, from the coincidence with the inertial frame to the actual orientation, and \mathbf{b} is the position vector of the body-frame origin, B, measured in the inertial frame.

$$\langle \dot{\mathbf{X}}_1, \dot{\mathbf{X}}_2 \rangle \triangleq \dot{\mathbf{X}}_1^T \mathbf{Q} \dot{\mathbf{X}}_2 \quad (1),$$

if a symmetric, positive-definite, 6×6 matrix, \mathbf{Q} , existed which makes $\dot{\mathbf{X}}_1^T \mathbf{Q} \dot{\mathbf{X}}_2$ independent of the reference frame (i.e., inertial or base frame), used to represent $\dot{\mathbf{X}}$, and of \mathbf{X} (i.e., of \mathbf{R} and \mathbf{b} which vary when either the inertial frame or the body frame is changed). Such a \mathbf{Q} matrix does not exist (see [4] for demonstration and references); whereas, symmetric, positive-definite \mathbf{Q} matrices that are independent of \mathbf{X} (i.e., of the inertial-frame and of the body-frame choices) exist, and, by using them, left-invariant and right-invariant distance metrics can be defined according to the representation of $\dot{\mathbf{X}}$ used in relationship (1).

Among the \mathbf{Q} matrices that generates left-invariant (right-invariant) metrics through definition (1), the following one has been suggested in [4]:

$$\mathbf{Q} = \begin{bmatrix} c_1 \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & c_2 \mathbf{I}_3 \end{bmatrix} \quad (2),$$

where c_1 and c_2 are two positive scalar constants, whereas $\mathbf{0}_3$ and \mathbf{I}_3 are the null and the identity 3×3 matrices, respectively. The left-invariant distance metric generated by using (2) and the inner product (1) is (see [4] for details):

$$\rho_P(\mathbf{X}_1, \mathbf{X}_2) = \sqrt{c_1 \delta_S^2(\mathbf{R}_1, \mathbf{R}_2) + c_2 \delta_T^2(\mathbf{b}_1, \mathbf{b}_2)} \quad (3)$$

where \mathbf{R}_i and \mathbf{b}_i are the rotation matrix and the position vector, respectively, which locate the i -th rigid-body pose and the corresponding SE(3) element, \mathbf{X}_i ($i=1,2$); whereas δ_S and δ_T are the following bi-invariant distance metrics of SO(3) and \mathbf{R}^3 , respectively (the image of $\cos^{-1}(\cdot)$ is restricted to the range $[0, \pi]$ radians, and $|\cdot|$ denotes the Euclidean norm of \mathbf{R}^3):

$$\delta_S(\mathbf{R}_1, \mathbf{R}_2) = \cos^{-1}\left(\frac{\text{tr}(\mathbf{R}_1^T \mathbf{R}_2) - 1}{2}\right), \quad (4a)$$

$$\delta_T(\mathbf{b}_1, \mathbf{b}_2) = |\mathbf{b}_2 - \mathbf{b}_1|. \quad (4b)$$

The c_1 and c_2 parameters and relationship (3) define a two-parameters family of left-invariant scale-dependent distance metrics. The geometric meaning of a limitation on $\rho_P(\mathbf{X}_1, \mathbf{X}_2)$ (i.e. an inequality of type $\rho_P(\mathbf{X}_1, \mathbf{X}_2) < \varepsilon$) is difficult to be found, while it is clear that c_1 and c_2 assign different weights to the rotational and translational parts of the geodesic path that brings the rigid body from \mathbf{X}_1 to \mathbf{X}_2 . Analogous considerations can be done on the right-invariant distance metrics generated by using (2) and the inner product (1).

The above-reported technique has the advantage of providing definitions which allow the length of rigid-body paths to be measured. Such an appealing feature is not required in many technical applications where only finitely separated poses have to be considered. In these cases, the direct introduction of a distance metric into SE(3) can be done without considering its tangent space. This was done, for instance, by Kazerounian and Rastegar in [2] where they proposed the following family of distance metrics, called object norms:

$$\rho_{KR}(\mathbf{X}_1, \mathbf{X}_2) = \frac{1}{V} \int_V w(\mathbf{P}) |\mathbf{P}(\mathbf{X}_2) - \mathbf{P}(\mathbf{X}_1)|^2 dV \quad (5)$$

where V is the volume of the rigid body that assumes the two poses identified by \mathbf{X}_1 and \mathbf{X}_2 ; \mathbf{P} is the position vector, measured in the inertial frame, that locates a generic point P of the rigid body; and $w(\mathbf{P})$ is a weighting function whose values depend on P .

Distance metrics assigned through relationship (5) measure weighted average square errors on the positioning of a given rigid body. They are bi-invariant, even though they depend on the rigid-body shape and are cumbersome to be repeatedly evaluated, for instance in optimization or control procedures, since their evaluation requires the computation of volume integrals. Relationship (5) allows the distance metric to be calibrated so that it matches suitable design criteria by changing the weighting function.

Other techniques can be used to directly introduce a distance metric. Among these techniques, it is worth mentioning the direct introduction of distance metrics by the definition of a mapping (kinematic mapping) [18], from the rigid-body c -space into a suitable image projective space, and the successive introduction of a distance metric into the image space of the kinematic mapping. Even though this technique, proposed by Ravani and Roth in [1] for planar motions and successively extended in [19, 20], gives the bases for the indirect introduction of distance metrics, it is a direct introduction since, in principle, the image space is just a different representation of the c -space (i.e., the kinematic mapping practically is a change of coordinates in the c -space) and the idea of

approximating an actual displacement with a fictitious one is not present here. In [1], the pose of a rigid body constrained to perform planar motion is given through four parameters, say q_i for $i=1, \dots, 4$, constrained to satisfy the following scalar equation

$$q_3^2 + q_4^2 = 1 \quad (6)$$

The quadruplet $\mathbf{q} \triangleq \{q_1, q_2, q_3, q_4\}^T$, named planar quaternion, identifies a point that lies on an hyper-surface, σ , of \mathbb{R}^4 . Such a hyper-surface is the geometric representation of the projective image space into which the c-space is mapped. With reference to the triplet of c-space coordinates $\{b_x, b_y, \phi\}$, where $\mathbf{b} = \{b_x, b_y\}^T$ is the position vector locating the origin B of the body frame in the inertial frame, and ϕ is the rotation angle the inertial frame must counterclockwise rotate to have the same orientation of the body frame (see Fig. 2), the kinematic mapping from the c-space into σ is given by the following relationships:

$$q_1 = \frac{1}{2} \left[b_x \sin\left(\frac{\phi}{2}\right) - b_y \cos\left(\frac{\phi}{2}\right) \right] \quad (7a)$$

$$q_2 = \frac{1}{2} \left[b_x \sin\left(\frac{\phi}{2}\right) + b_y \cos\left(\frac{\phi}{2}\right) \right] \quad (7b)$$

$$q_3 = \sin\left(\frac{\phi}{2}\right) \quad (7c)$$

$$q_4 = \cos\left(\frac{\phi}{2}\right) \quad (7d)$$

In the image space, the following distance metric is defined [1]:

$$\rho_{RR}(\mathbf{q}_1, \mathbf{q}_2) = \sqrt{(\mathbf{q}_2 - \mathbf{q}_1)^T (\mathbf{q}_2 - \mathbf{q}_1)} \quad (8)$$

where \mathbf{q}_1 and \mathbf{q}_2 are the two planar quaternions corresponding to the two poses whose distance has to be computed.

Relationship (8) is the Euclidean norm of \mathbb{R}^4 , it is scale-dependent and is not bi-invariant [3]. The design constraint given through a limitation on $\rho_{RR}(\mathbf{q}_1, \mathbf{q}_2)$ (i.e., an inequality of type $\rho_{RR}(\mathbf{q}_1, \mathbf{q}_2) < \varepsilon$) is difficult to be identified.

3. INDIRECT INTRODUCTION OF DISTANCE METRICS IN SE(3)

McCarthy, in [9], showed how planar and spatial rigid-body motions can be approximated by spherical and three-spherical motions, respectively.

He exploited the fact that the motion of an oriented great-circle arc on a sphere with radius r completely identifies a rigid-body spherical motion centered at the sphere center, and that, if r is sufficiently large, the spherical motion approximately describe a planar motion in a plane tangent to the sphere. By formalizing this concept, he stated a first-order

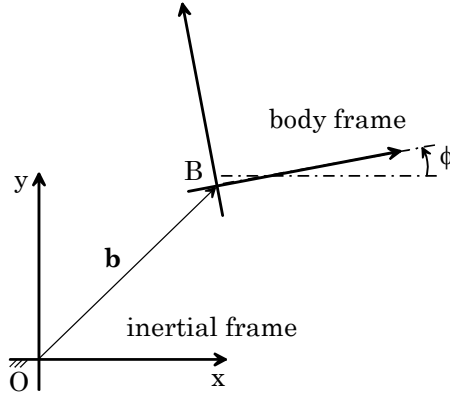


Figure 2. Inertial frame and body frame in the case of planar motion.

correspondence in r^{-1} between the 3×3 homogeneous transformation matrix, which identifies the pose of a rigid body constrained to perform planar motion, and a rotation matrix of $SO(3)$. Then, he generalized this approximation technique till to state a first-order correspondence in r^{-1} between the 4×4 homogeneous transformation matrix, which identifies the finite motion that brings the body frame from the coincidence with the inertial frame to its actual pose, and a 4×4 orthogonal matrix with determinant equal to 1, which is an element of $SO(4)$ and identifies a six-parameters non-physical motion, named three-spherical motion. These correspondences can be seen as kinematic mappings, from the c-space into $SO(3)$ or $SO(4)$, which depend on the parameter r^{-1} . Once they are stated the distance metrics of the image space (i.e. $SO(3)$ or $SO(4)$) can be adopted as distance metrics of the c-space [11], same as it was done in [1]. In the remaining part of this section, this procedure will be briefly illustrated.

The sixteen entries of a 4×4 orthogonal matrix satisfy ten scalar conditions (i.e., four conditions to impose that the columns are unit vectors, plus 6 conditions to impose that the four columns are mutually orthogonal), thus they can be parameterized as a function of only six parameters. The four column vectors of such a matrix constitute a basis of R^4 , and they can be considered quadru-dimensional vectors that identify the directions of four coordinate axes, say x, y, z, and w, that, when combined two by two, locate six coordinate hyper-planes: xy, yz, zx, xw, yw, zw. An elementary rotation in R^4 can be defined as a quadru-dimensional motion of a quadru-dimensional vector, with constant magnitude, that changes only two components of that vector measured in the reference basis. The quadru-dimensional vector collecting the new components after the rotation can be obtained by multiplying a suitable 4×4 orthogonal matrix by the quadru-dimensional vector collecting the old component (i.e., the ones before the rotation). Such suitable matrices will be called elementary rotation matrices. Since six couples of coordinate axes can be counted in R^4 , as many elementary rotations (one for each coordinate hyper-plane) can be defined in R^4 , and the associated elementary rotation matrices are ($c_{(\cdot)}$ and $s_{(\cdot)}$ denote $\cos(\cdot)$ and $\sin(\cdot)$, respectively)

$$\mathbf{H}_{xy}(\phi) = \begin{bmatrix} c_\phi & -s_\phi & 0 & 0 \\ s_\phi & c_\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{H}_{zx}(\psi) = \begin{bmatrix} c_\psi & 0 & s_\psi & 0 \\ 0 & 1 & 0 & 0 \\ -s_\psi & 0 & c_\psi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{H}_{yz}(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_\theta & -s_\theta & 0 \\ 0 & s_\theta & c_\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad (9a)$$

$$\mathbf{H}_{xw}(\alpha) = \begin{bmatrix} c_\alpha & 0 & 0 & s_\alpha \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -s_\alpha & 0 & 0 & c_\alpha \end{bmatrix}; \mathbf{H}_{yw}(\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_\beta & 0 & s_\beta \\ 0 & 0 & 1 & 0 \\ 0 & -s_\beta & 0 & c_\beta \end{bmatrix}; \mathbf{H}_{zw}(\gamma) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_\gamma & s_\gamma \\ 0 & 0 & -s_\gamma & c_\gamma \end{bmatrix}; \quad (9b)$$

Moreover, the generic element, \mathbf{H} , of $\text{SO}(4)$ can always be written as follows:

$$\mathbf{H} = \mathbf{J}(\alpha, \beta, \gamma) \mathbf{K}(\theta, \psi, \phi) \quad (10)$$

with

$$\mathbf{J}(\alpha, \beta, \gamma) = \mathbf{H}_{zw}(\gamma) \mathbf{H}_{yw}(\beta) \mathbf{H}_{xw}(\alpha) = \begin{bmatrix} c_\alpha & 0 & 0 & s_\alpha \\ -s_\alpha s_\beta & c_\beta & 0 & c_\alpha s_\beta \\ -s_\alpha c_\beta s_\gamma & -s_\beta s_\gamma & c_\gamma & c_\alpha c_\beta s_\gamma \\ -s_\alpha c_\beta c_\gamma & -s_\beta c_\gamma & -s_\gamma & c_\alpha c_\beta c_\gamma \end{bmatrix} \quad (11a)$$

$$\mathbf{K}(\theta, \psi, \phi) = \mathbf{H}_{xy}(\phi) \mathbf{H}_{zx}(\psi) \mathbf{H}_{yz}(\theta) = \begin{bmatrix} \mathbf{R}(\theta, \psi, \phi) & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (11b)$$

where $\mathbf{R}(\theta, \psi, \phi)$ is an element of $\text{SO}(3)$ (i.e., a 3×3 rotation matrix) parameterized through the angles θ , ψ , and ϕ that are the Roll, Pitch, and Yaw angles of the RPY parameterization [21]. It is worth noting that the subset $\{\mathbf{H} \in \text{SO}(4) \mid \alpha = \beta = \gamma = 0\}$ of $\text{SO}(4)$ is a different representation of $\text{SO}(3)$ made through 4×4 matrices (i.e., with this representation $\text{SO}(3)$ is a subgroup of $\text{SO}(4)$).

The generic element, $\mathbf{X} \triangleq \begin{bmatrix} \mathbf{R} & \mathbf{b} \\ \mathbf{0}^\top & 1 \end{bmatrix}$, of $\text{SE}(3)$ can be mapped into an element $\mathbf{H}(\mathbf{b}, \mathbf{R})$ of $\text{SO}(4)$ through the following analytic relationships ($\mathbf{b} = \{b_x, b_y, b_z\}^\top$):

$$\alpha = \frac{b_x}{r}; \quad \beta = \frac{b_y}{r}; \quad \gamma = \frac{b_z}{r}; \quad (12)$$

where r is a parameter denoting the radius of an hyper-sphere of \mathbf{R}^4 . The coordinates of a generic point, P , lying on this hyper-sphere can be given through the following quadru-dimensional position vector

$$\hat{\mathbf{P}} = r \{s_\xi c_\eta c_\zeta, s_\eta c_\zeta, s_\zeta, c_\xi c_\eta c_\zeta\}^\top \quad (13)$$

where ξ , η , and ζ are three angular parameters that locate the point on the quadru-dimensional hyper-sphere. It is worth noting that $\hat{\mathbf{P}}^T \hat{\mathbf{P}}$ is always equal to r^2 (i.e., the equation of the hyper-sphere, $\hat{\mathbf{P}}^T \hat{\mathbf{P}} = r^2$, is satisfied).

A position vector, $\mathbf{p} = \{p_x, p_y, p_z\}^T$, which locates a point P of \mathbb{R}^3 , can be mapped into a quadru-dimensional position vector, $\hat{\mathbf{P}}$, which locates a point of the hyper-sphere, through the following analytic relationships:

$$\xi = \frac{p_x}{r}; \quad \eta = \frac{p_y}{r}; \quad \zeta = \frac{p_z}{r}; \quad (14)$$

If the radius, r , of the hyper-sphere is sufficiently greater both than the components of \mathbf{b} and than the components of \mathbf{p} , so that the angles α , β , γ , ξ , η , and ζ , given by the linear mappings (13) and (14), are small enough to approximate their sine with the argument and their cosine with 1, the following relationship hold:

$$\mathbf{H} = \begin{bmatrix} \mathbf{R} & \frac{1}{r} \mathbf{b} \\ \mathbf{0}^T & 1 \end{bmatrix}; \quad \hat{\mathbf{P}} = \begin{Bmatrix} \mathbf{p} \\ r \end{Bmatrix}. \quad (15)$$

The analysis of (15) shows that, by increasing r , the hyper-sphere points can be confused with the points of the hyper-plane $w = r$, which give a particular homogeneous representation of the points of \mathbb{R}^3 . Moreover, it reveals that the product $\mathbf{H} \hat{\mathbf{P}}$ is exactly equal to $\{(\mathbf{R} \mathbf{p} + \mathbf{b})^T, r\}^T$, which gives the homogeneous coordinates of the point P, located by \mathbf{p} in \mathbb{R}^3 after the displacement, \mathbf{X} of SE(3), given by \mathbf{R} and \mathbf{b} , so matrix \mathbf{H} can be considered another representation of the element \mathbf{X} of SE(3). Therefore, the mapping of SE(3) into SO(4) given by (12), together with (13) and (14), identifies a correspondence between 4x4 matrices that tend to coincide with one another when r increases.

An one-to-one correspondence between the elements of SO(4) and suitable couples of unit quaternions, named biquaternions, has been presented in [22]. By exploiting this correspondence, the following distance metric of SO(4), which becomes a distance metric of SE(3) through relationships (11) and (12), has been proposed in [11]:

$$\rho_{EM}(\mathbf{H}_1, \mathbf{H}_2) = \sqrt{(\hat{\mathbf{g}}_2 - \hat{\mathbf{g}}_1)(\hat{\mathbf{g}}_2 - \hat{\mathbf{g}}_1)^* + (\hat{\mathbf{h}}_2 - \hat{\mathbf{h}}_1)(\hat{\mathbf{h}}_2 - \hat{\mathbf{h}}_1)^*} \quad (16)$$

where \mathbf{H}_1 and \mathbf{H}_2 are the two elements of SO(4) whose distance has to be measured, $(\hat{\mathbf{g}}_i, \hat{\mathbf{h}}_i)$ for $i=1,2$ is the couple of unit quaternions corresponding to \mathbf{H}_i , and $(\cdot)^*$ denotes the conjugate quaternion of (\cdot) .

Relationship (16), proposed in [11], defines a distance metric that is bi-invariant. Another bi-invariant distance metric of SO(4) can be defined as follows [12]:

$$\rho_L(\mathbf{H}_1, \mathbf{H}_2) = \left\| \mathbf{I}_4 - \mathbf{H}_2 \mathbf{H}_1^T \right\|_F \quad (17)$$

where \mathbf{I}_4 is the 4×4 identity matrix, and $\|\mathbf{A}\|_F$ denotes the Frobenius norm of matrix \mathbf{A} .

4. GENERATION OF SE(3) DISTANCE METRICS

This author, in [17], has proposed a technique for generating new distance metrics of SE(3) by using already defined distance metrics that can be either directly or indirectly introduced in SE(3). This technique has the advantage of introducing arbitrary coefficients, into the final expression of the distance metric, which can be calculated to match design criteria. In this section this technique will be briefly recalled.

A distance metric in an n-dimensional c-space ($n \leq 6$) is a real-valued mapping $d: \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R}$ of type $d(\mathbf{y}_1, \mathbf{y}_2)$, where $\mathbf{y}_1, \mathbf{y}_2 \in \mathbf{R}^n$ identify two poses in the c-space, that satisfies the following properties

$$\begin{aligned} d(\mathbf{y}_1, \mathbf{y}_2) &= d(\mathbf{y}_2, \mathbf{y}_1); \text{ (simmetry)} \\ d(\mathbf{y}_1, \mathbf{y}_2) &> 0 \text{ if } \mathbf{y}_1 \neq \mathbf{y}_2 \text{ and } d(\mathbf{y}_1, \mathbf{y}_2) = 0 \text{ if } \mathbf{y}_1 = \mathbf{y}_2; \text{ (positive definiteness)} \\ d(\mathbf{y}_1, \mathbf{y}_2) &\leq d(\mathbf{y}_1, \mathbf{y}_3) + d(\mathbf{y}_3, \mathbf{y}_2) \text{ for any } \mathbf{y}_3 \in \mathbf{R}^n; \text{ (triangle inequality)} \end{aligned}$$

It is easy to demonstrate the following proposition:

Proposition 1: If $d_i(\mathbf{y}_1, \mathbf{y}_2)$ for $i=1, \dots, m$ are m different distance metrics defined on the same n-dimensional c-space, then any linear combination, $\delta(\mathbf{y}_1, \mathbf{y}_2)$, with non-negative coefficients of the $d_i(\mathbf{y}_1, \mathbf{y}_2)$ is a distance metric of the same c-space, provided that at least one coefficient be different from zero (in formulae: $\delta(\mathbf{y}_1, \mathbf{y}_2) \in \{ \sum_{i=1, m} a_i d_i(\mathbf{y}_1, \mathbf{y}_2) \mid (a_i \geq 0)$

$\&(\exists k : a_k \neq 0)\}$).

Considering the displacement of a body frame from the coincidence with the inertial frame to an assigned rigid-body pose, and, then, giving the geometric parameters of that displacement is a common way to identify the assigned pose in the c-space.

The special Euclidean group, SE(3), collects all the possible displacements of a rigid body. SE(3) admits 10 subgroups [23] with dimension greater than zero and lower than six (lower-mobility subgroups).

A generic element of SE(3) can be always obtained through the composition of elements that belong to particular lower-mobility subgroups (e.g. one translation composed with one spherical motion). This property allows a rigid-body pose to be parameterized through a six-tuple, \mathbf{x} , whose entries can be collected into p n_k -tuples, ${}^k\mathbf{y}$ with $k=1, \dots, p$ and $\sum_{k=1, p} n_k = 6$, where each ${}^k\mathbf{y}$ identifies an element of a given lower-mobility subgroup, S_k

with dimension n_k (in formulae: $\mathbf{x} \in \{ ({}^1\mathbf{y}^T, \dots, {}^p\mathbf{y}^T)^T \mid {}^k\mathbf{y} \in S_k \subset SE(3) \}$ ⁶). In this context, it is easy to demonstrate the following proposition:

Proposition 2: If $\delta_k({}^k\mathbf{y}_1, {}^k\mathbf{y}_2)$ is a distance metric of S_k , then any linear combination, $\rho(\mathbf{x}_1, \mathbf{x}_2)$, where $\mathbf{x}_1 = ({}^1\mathbf{y}_1^T, \dots, {}^p\mathbf{y}_1^T)^T$ and $\mathbf{x}_2 = ({}^1\mathbf{y}_2^T, \dots, {}^p\mathbf{y}_2^T)^T$, of the $\delta_k({}^k\mathbf{y}_1, {}^k\mathbf{y}_2)$ with positive coefficients is a distance metric of SE(3) (in formulae: $\rho(\mathbf{x}_1, \mathbf{x}_2) \in \{ \sum_{k=1, p} b_k \delta_k({}^k\mathbf{y}_1, {}^k\mathbf{y}_2) \mid$

⁶ The elements of SE(3) are 4×4 transformation matrices that depend on six geometric parameters. These parameters are the ones collected in the six-tuple \mathbf{x} . In this section, for simplifying the notation, the six-tuple \mathbf{x} is confused with the element of SE(3) computed through the entries of \mathbf{x} . The same notation simplification is used for the subgroups of SE(3).

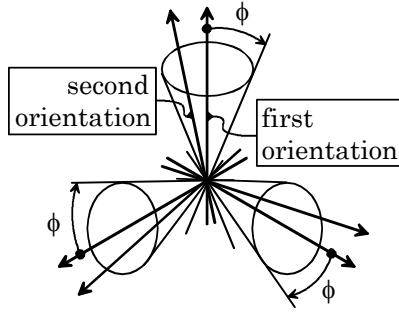


Figure 3. Geometric meaning of the condition $\delta_S(\mathbf{R}_1, \mathbf{R}_2) < \phi$.

$$(\mathbf{x}_1 = ({}^1\mathbf{y}_1^T, \dots, {}^p\mathbf{y}_1^T)^T) \& (\mathbf{x}_2 = ({}^1\mathbf{y}_2^T, \dots, {}^p\mathbf{y}_2^T)^T) \& (\forall b_k > 0)$$

Proposition (2) makes it possible to generate a big family of distance metrics of $SE(3)$ by decomposing a generic displacement into displacements of lower-mobility subgroups and by defining one distance metric in each subgroup. Once a family of distance metrics is determined, how to select meaningful distance metrics among the members of this family still is an open problem. In the next subsection this problem will be addressed by looking for distance metrics that have an immediate geometric meaning.

Example of Application

The decomposition of a generic rigid-body displacement into displacements that belong to lower-mobility subgroups can be implemented in many ways. Nevertheless, only some subgroups have distance metrics that are easy to use and with a straightforward geometric interpretation. The subgroup of the spatial translations, $T(3)$, and the subgroup of the spherical motions, $S(3)$, are among these subgroups. Since any displacement can be obtained by composing one spatial translation with one spherical motion, $T(3)$ and $S(3)$ will be used to decompose spatial displacements.

When a rigid body is constrained to translate, its pose (\equiv position) is uniquely identified by the coordinates of the origin, B (Fig. 1), of the body frame measured in the inertial frame. The distance metric $\delta_T(\mathbf{b}_1, \mathbf{b}_2)$ defined by relationship (4b) is commonly adopted in $T(3)$. $\delta_T(\mathbf{b}_1, \mathbf{b}_2)$ is bi-invariant in $T(3)$. Moreover, a limitation on $\delta_T(\mathbf{b}_1, \mathbf{b}_2)$ (e.g. $\delta_T(\mathbf{b}_1, \mathbf{b}_2) < c$) has a clear geometric meaning. In fact, it means that the position B_2 of B must be located inside a sphere with center at the position B_1 of B and radius given by the imposed condition.

When a rigid body is constrained to perform spherical motions with the same center, hereafter assumed coincident with the origin of the body frame, its pose (\equiv orientation) is uniquely identified by the rotation matrix \mathbf{R} , whose column are the three unit vectors of the body-frame axes projected onto the inertial frame (Fig. 1). The set that collects all the rotation matrices is named $SO(3)$, and the above considerations state an isomorphism between $S(3)$ and $SO(3)$. The distance metric $\delta_S(\mathbf{R}_1, \mathbf{R}_2)$, defined by relationship (4a), can be adopted in $SO(3)$. $\delta_S(\mathbf{R}_1, \mathbf{R}_2)$ is bi-invariant in $SO(3)$. It measures the convex rotation angle, in radian, of the finite rotation that brings the first rigid-body orientation to coincide with the second one. On a unit sphere centered at the center of the spherical motion, such

an angle measures the length of the great-circle arc between two points, that lie on the equatorial circle perpendicular to the rotation axis and coincide with each other after the above-mentioned finite rotation. A limitation on $\delta_S(\mathbf{R}_1, \mathbf{R}_2)$ (e.g. $\delta_S(\mathbf{R}_1, \mathbf{R}_2) < \phi$) has a clear geometric meaning. In fact, it means that each body-frame axis at the second orientation is confined to lie inside a circular cone with vertex at the center of the spherical motion, cone axis coincident with the homologous body-frame axis at the first orientation, and vertex angle given by the imposed condition (Fig. 3).

The distance metrics (4a) and (4b), and proposition (2) can be used to generate the following family of distance metrics of SE(3):

$$\rho_u(\mathbf{X}_1, \mathbf{X}_2) = \delta_T(\mathbf{b}_1, \mathbf{b}_2) + u \delta_S(\mathbf{R}_1, \mathbf{R}_2) \quad (18)$$

where u is an arbitrary positive constant that is evaluated in the same unit as δ_T .

The analysis of definition (3) reveals that a limitation on $\rho_u(\mathbf{X}_1, \mathbf{X}_2)$ expressed in the following form

$$\rho_u(\mathbf{X}_1, \mathbf{X}_2) < h \quad (19)$$

implies the following limitations on δ_T and δ_S , and the associated geometric meanings

$$\delta_T(\mathbf{b}_1, \mathbf{b}_2) < h \quad (20a)$$

$$\delta_S(\mathbf{R}_1, \mathbf{R}_2) < \frac{h}{u} \quad (20b)$$

Since the value of $\rho_u(\mathbf{X}_1, \mathbf{X}_2)$ depends on the choice of the origin, B , of the body frame, the distance metrics defined by (18) are, in general, left-invariant.

5. CONCLUSIONS

A critical review of the most known distance metrics of SE(3), proposed in the literature, has been presented.

In particular, both the direct and the indirect distance-metric introductions into SE(3) have been considered together with a technique, proposed by the author [17], for generating new distance metrics by combining already defined distance metrics that can be either directly or indirectly introduced in SE(3). This technique has the advantage of introducing arbitrary coefficients, into the final expression of the distance metric, which can be calculated to match design criteria.

Acknowledgements

This work has been supported by MIUR funds.

The author wishes to thank the GEF 2008 organizers for having charged on their funds the costs of the event.

REFERENCES

- [1] Ravani, B., and Roth, B., 1983, "Motion synthesis using kinematic mappings," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 105, pp. 460-467.
- [2] Kazerounian, K., and Rastegar, J. 1992, "Object norms: a class of coordinate and metric independent norms for displacements," *Proceedings of the ASME 1992 Design Engineering Technical Conferences*, Scotsdale, USA.
- [3] Rico Martinez, J.M., and Duffy, J., 1995, "On the metrics of rigid body displacements for infinite and finite bodies," *ASME Journal of Mechanical Design*, Vol. 117, No. 1, pp. 41-47.
- [4] Park, F.C., 1995, "Distance metrics on the rigid-body motion with application to mechanism design," *ASME Journal of Mechanical Design*, Vol. 117, No. 1, pp. 48-54.
- [5] Zefran, M., Kumar, V., and Croke, C., 1996, "Choice of Riemannian metrics for rigid body kinematics," *Proceedings of the ASME 1996 Design Engineering Technical Conferences*, Irvine, USA.
- [6] Lin, Q., and Burdick, J., 2000, "Objective and frame-invariant kinematic metric functions for rigid bodies," *The International Journal of Robotics Research*, Vol. 19, No. 6, pp. 612-625.
- [7] Eberharter, J., and Ravani, B., 2004, "Local metrics for rigid body displacements," *ASME Journal of Mechanical Design*, Vol. 126, No. 5, pp. 805-812.
- [8] Zhang, Y., and Ting, K.-L., 2005, "Point-line distance under Riemannian metrics," *Proceedings of the ASME 2005 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Long Beach, USA.
- [9] McCarthy, J.M., 1983, "Planar and spatial rigid body motion as special cases of spherical and 3-spherical motion," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 105, pp. 569-575.
- [10] Larochelle, P., and McCarthy, J.M., 1995, "Planar motion synthesis using an approximate bi-invariant metric," *ASME Journal of Mechanical Design*, Vol. 117, No. 4, pp. 646-651.
- [11] Etzel, K., and McCarthy, J.M., 1996, "A metric for spatial displacements using biquaternions on SO(4)," *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, Minneapolis, USA.
- [12] Larochelle, P., 1999, "On the geometry of approximate bi-invariant projective displacement metrics," *Proceedings of the 10th World Congress on the Theory of Machines and Mechanisms*, Oulu, Finland, Vol. 2, pp. 548-553.
- [13] Tse, D.M., and Larochelle, P.M., 2000, "Approximating spatial locations with spherical orientations for spherical mechanism design," *ASME Journal of Mechanical Design*, Vol. 122, No. 4, pp. 457-463.
- [14] Belta, C., and Kumar, V., 2002, "An SVD-based projection method for interpolation on SE(3)," *IEEE Transactions on Robotics and Automation*, Vol. 18, No. 3, pp. 334-345.
- [15] Angeles, J., 2005, "Is there a characteristic length of a rigid-body displacement?," *Proceedings of the 2005 International Workshop on Computational Kinematics*, Cassino, Italy.
- [16] Larochelle, P., Murray, A.P., and Angeles, J., 2007, "A distance metric for finite sets of rigid-body displacements via the polar decomposition," *ASME Journal of Mechanical Design*, Vol. 129, No. 8, pp. 883-886.

- [17] Di Gregorio R., 2008, "A novel point of view to define the distance between two rigid-body poses," *Advances in Robot Kinematics: Analysis and Design*, ARK2008, Eds. J. Lenarcic and P. Wenger, Springer Science + Business Media Publishers, Amsterdam (the Netherlands), ISBN: 978-1-4020-8599-4, pp. 361-369.
- [18] Bottema, O., and Roth, B., 1979, "Theoretical kinematics," North-Holland Publishers, Amsterdam (the Netherlands).
- [19] Ravani, B., and Roth, B., 1984, "Mappings of spatial kinematics," *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, Vol. 106, pp. 341-347.
- [20] Eberharder, J.K., and Ravani, B., 2004, Local metrics for rigid body displacements," *ASME Journal of Mechanical Design*, Vol. 126, pp. 805-812.
- [21] Tsai, L.-W., 1999, "Robot analysis: the mechanics of serial and parallel manipulators," John Wiley & Sons, Inc., ISBN: 0-471-32593-7, pp. 38-39.
- [22] Ge, Q.J., 1998, "On the matrix realization of the theory of biquaternion," *ASME Journal of Mechanical Design*, Vol. 120, No. 3, pp. 404-407.
- [23] Hervé, J.M., 1978, "Analyse structurale des mécanismes par groupe des déplacements," *Mechanism and Machine Theory*, Vol. 13, No. 4, pp. 437-450.

DIMENSIONAMENTO MECCANICO, CONTROLLO E PROTOTIPAZIONE DI ATTUATORI A MEMORIA DI FORMA

Eugenio Dragoni
*Dipartimento di Scienze e Metodi dell'Ingegneria,
Università di Modena e Reggio Emilia, Italy
E-mail: eugenio.dragoni@unimore.it*

Giovanni Scirè Mammano
*Dipartimento di Scienze e Metodi dell'Ingegneria,
Università di Modena e Reggio Emilia, Italy
E-mail: giovanni.sciremammano@unimore.it*

Igor Spinella
*Dipartimento di Scienze e Metodi dell'Ingegneria,
Università di Modena e Reggio Emilia, Italy
E-mail: igor.spinella@unimore.it*

Abstract. *Il lavoro è incentrato sul dimensionamento meccanico, sul controllo e sulla prototipazione di attuatori a memoria di forma (SMA). I materiali a memoria di forma sono leghe metalliche in grado di recuperare una forma memorizzata se riscaldati. Tali materiali hanno molteplici campi di applicazione, tra i quali il loro utilizzo come attuatori. Nel presente lavoro vengono sviluppate formule utili al dimensionamento ottimale di attuatori a memoria di forma, si propone una metodologia di controllo economica ed efficace di tali attuatori e si illustra un prototipo di attuatore SMA, utilizzato per realizzare una valvola oleoidraulica di interesse industriale.*

Keywords: *SMA, Shape Memory Alloy, Actuator Design, Position Control, Actuator Prototyping*

1. INTRODUZIONE

Cosa sono i materiali a memoria di forma

I materiali a memoria di forma (Shape Memory Alloys – SMA) sono leghe metalliche binarie o ternarie capaci di recuperare una forma memorizzata se riscaldate al di sopra di una certa temperatura critica. Tale proprietà è detta memoria di forma (Shape Memory Effect – SME).

Oltre alla memoria di forma, gli SMA mostrano effetti quali la pseudo elasticità (detta anche superelasticità), la pseudoplasticità e l'isteresi nei cicli di funzionamento. Gli effetti di memoria di forma, di pseudoelasticità e di pseudoplasticità [1] sono legati a trasformazioni di fase del reticolo cristallino indotte da variazioni di temperatura (memoria di forma) e da variazioni di tensione meccanica (pseudoelasticità e pseudoelasticità). Il

fatto che le trasformazioni non siano termodinamicamente reversibili determina isteresi nel comportamento dei materiali a memoria di forma stessi.

La Fig. 1 mostra il tipico comportamento di una lega a memoria di forma sottoposta a trazione a diverse temperature [2]. In blu è riportato il comportamento a bassa temperatura, mentre in rosso il comportamento ad alta temperatura. A bassa temperatura e bassa tensione meccanica è stabile una fase detta martensite twinned, ovvero senza orientazioni preferenziali nello spazio e caratterizzata da una struttura cristallina monoclina. Sottoponendo a trazione la martensite twinned (curva blu), dopo un primo tratto elastico, caratterizzato da un basso modulo elastico, si ha il fenomeno della pseudoplasticità: il comportamento apparente del materiale è plastico, ma la deformazione è in realtà legata ad un progressivo riorientamento della martensite twinned nella direzione dello sforzo. La martensite orientata è detta martensite detwinned. Se il materiale viene riscaldato in assenza di tensione meccanica, la martensite si trasforma in una fase cubica a corpo centrato detta austenite e la lega recupera la forma memorizzata (memoria di forma). La trasformazione pseudoplastica è dunque reversibile, qualora sia applicato alla lega un opportuno stimolo termico.

Si consideri adesso la curva rossa, relativa alla trazione di una lega SMA ad alta temperatura. Il materiale è inizialmente in stato austenitico, dunque, dopo un primo tratto elastico, caratterizzato da un elevato modulo elastico E_a , si ha un plateau a tensione quasi costante dovuto alla trasformazione pseudoelastica. Nel passaggio da uno stato all'altro avviene una trasformazione di fase da austenite a martensite detwinned indotta da tensione meccanica. Qualora si scarichi progressivamente il materiale, la martensite detwinned si ritrasforma in austenite, ma a livelli di tensione meccanica più bassi. Ciò determina un curva caratterizzata da isteresi: l'area compresa tra le curve di caricamento e scaricamento rappresenta l'energia dissipata per unità di volume dalla lega a memoria di forma.

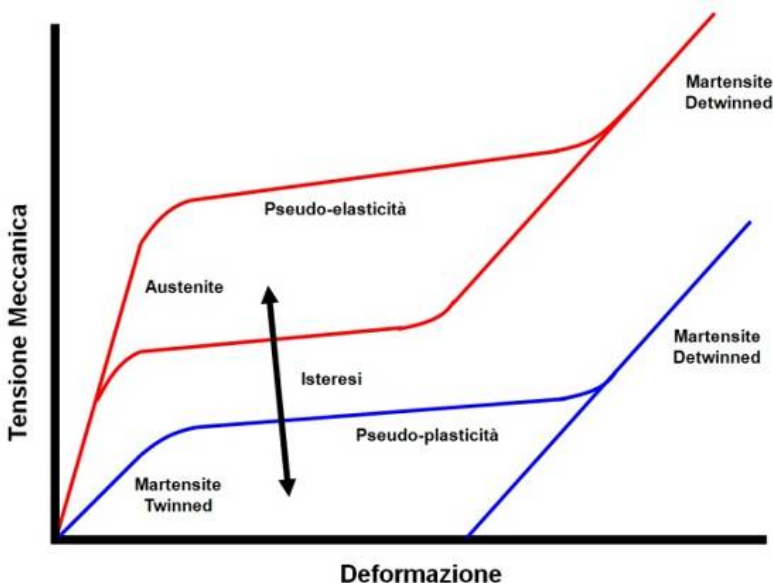


Figura 1. Comportamento delle leghe a memoria di forma

Principali leghe a memoria di forma

Le leghe a memoria di forma sono molteplici, ma le più diffuse e utilizzate sono le leghe ternarie di rame (CuAlBe, CuAlNi, CuZnAl) e le leghe binarie di Nickel – Titanio [3]. Queste ultime sono in assoluto le più sfruttate e sono conosciute anche col nome di leghe NiTiNOL (Nickel – Titanium Naval Ordnance Labs – dal nome dei laboratori dove furono scoperte, all’inizio degli anni ‘60). La grande diffusione delle leghe NiTiNOL è legata soprattutto alle ottime proprietà meccaniche, alla resistenza agli ambienti ostili ed alla biocompatibilità.

Principi di utilizzazione e campi di applicazione

Ciascuno dei fenomeni descritti, caratteristici dei materiali a memoria di forma, può essere impiegato utilmente in dispositivi di interesse industriale, commerciale o biomedicale. Il fenomeno in assoluto più sfruttato è quello della superelasticità: molteplici sono i dispositivi biomedicali sviluppati, quali stent aortocoronarici, filtri vascolari e altri sistemi espandibili [4], o fili per dispositivi ortodontici [5]. La superelasticità è sfruttata in questi casi per esercitare una forza costante ed ottimale sui tessuti biologici a prescindere dallo stato di deflessione del dispositivo. Ad esempio, uno stent aortocoronarico aderisce perfettamente all’arteria coronaria nel quale è inserito, mantenendola pervia e senza rischiare di danneggiarla, a prescindere dalle variazioni di diametro della stessa (Fig. 2(a)). Lo stesso principio viene sfruttato con efficacia per realizzare giunzioni e raccordi (Fig. 2(b)). Il giunto sarà di rapido montaggio, poiché in fase di montaggio il giunto è freddo, e dunque pseudoplastico, mentre una volta raggiunta la temperatura di esercizio il giunto recupererà la forma ed entrerà in campo superelastico. Dunque, lo stesso eserciterà la corretta forza sui membri da collegare, a prescindere da eventuali disallineamenti anche consistenti tra gli stessi. La superelasticità, inoltre, viene sfruttata per realizzare occhiali, antenne o altri oggetti capaci di non arrivare a rottura neppure se sottoposti a notevoli deformazioni (Fig. 2(c)).

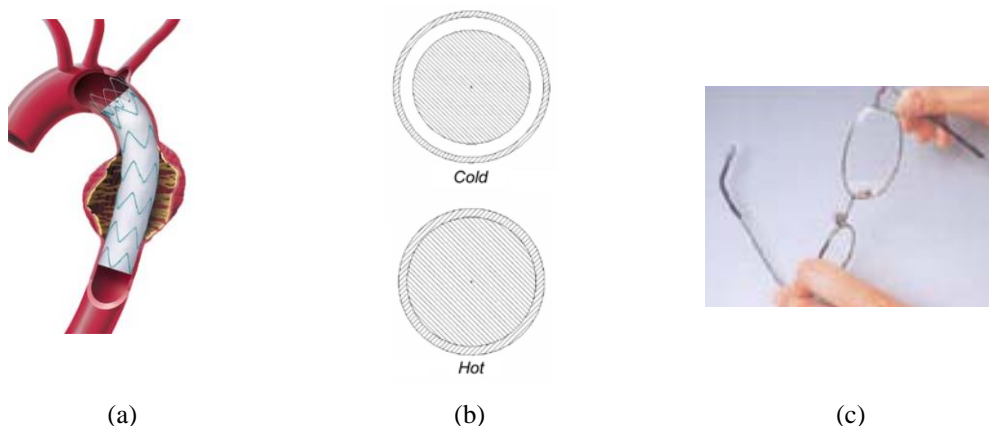


Figura 2. (a) Stent vascolare; (b) Giunto superelastico; (c) Occhiali a memoria di forma

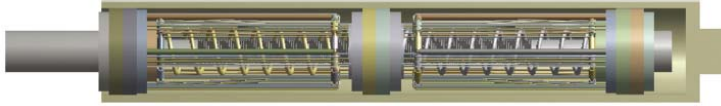


Figura 3. Smorzatore a memoria di forma

L'isteresi propria dei materiali a memoria di forma, invece, può essere sfruttata efficacemente per realizzare smorzatori allo stato solido [6], come quello di Fig. 3. Tale esempio di smorzatore è realizzato mediante due gruppi di fili superelastici pretensionati e contrapposti tra loro per garantire che l'isteresi propria del ciclo superelastico (curva rossa di Fig. 1) sia sfruttata per dissipare energia in entrambi i versi di movimentazione del sistema.

Altre applicazioni dell'isteresi sono individuabili nel campo del controllo delle vibrazioni [7], sia attivo che passivo: l'isteresi propria del materiale, infatti, è regolabile in funzione della temperatura alla quale la lega viene portata. Ciò consente, ad esempio, di realizzare molle o pannelli compositi o altre strutture in cui fili a memoria di forma sono adibiti al controllo attivo o passivo delle vibrazioni.

Infine, uno dei campi di applicazione della memoria di forma tra i più affascinanti e ancora non completamente esplorati è quello degli attuatori, campo sul quale è incentrato il presente lavoro.

Attuatori a memoria di forma

Come detto, un materiale a memoria di forma è in grado di recuperare una forma memorizzata precedentemente qualora venga riscaldato.

Le temperature critiche delle quali tener conto in un ciclo di attuazione sono quattro: A_s , temperatura di inizio trasformazione austenitica, A_f , temperatura di fine trasformazione austenitica, M_s , temperatura di inizio trasformazione martensitica, ed M_f , temperatura di fine trasformazione martensitica. Scaldando il materiale sopra ad A_s , il reticolo cristallino dello stesso inizia a trasformarsi da martensitico ad austenitico. Questa trasformazione si completa al raggiungimento di $A_f > A_s$. Dunque, al di sopra di A_f il materiale è completamente austenitico ed è caratterizzato dal modulo elastico dell'austenite, E_a . Se portato oltre una certa tensione meccanica critica, il materiale esibisce un comportamento pseudoelastico.

Viceversa, raffreddando il materiale, la trasformazione da austenite a martensite inizia ad una temperatura $M_s < A_s$. La trasformazione ha fine quando il materiale raggiunge la temperatura $M_f < M_s$. Dunque, sotto ad M_f il materiale è completamente martensitico ed è caratterizzato dal modulo elastico della martensite, E_m . Se portato oltre una certa tensione meccanica critica, il materiale esibisce un comportamento pseudoplastico.

Si ha che $E_a > E_m$, dunque il ciclo di riscaldamento / raffreddamento determina una variazione di modulo elastico che può essere sfruttata efficacemente per realizzare attuatori caratterizzati da elevatissime potenze specifiche rispetto ad ogni altra tipologia di attuatore, come mostrato in Fig. 4 [1].

D'altro canto, basso rendimento proprio delle leghe a memoria di forma e la ridotta bassa banda passante degli stessi, legata principalmente al difficile raffreddamento dei materiali stessi in tempi rapidi, riducono l'applicazione delle leghe stesse al campo della microattuazione, come mostrato da Fig. 4.

Oltre all'elevatissima potenza specifica, gli attuatori a memoria di forma sono silenziosi, puliti, biocompatibili e possono di lavorare in ambienti ostili. Infine, possono fungere al contempo da attuatore e da sensore, permettendo di realizzare sistemi integrati [1]. Infatti, oltre alla variazione di modulo elastico che risulta utile per l'attuazione, le trasformazioni proprie degli SMA determinano anche variazioni delle proprietà elettriche della leghe stesse (resistenza e induttanza) che possono essere sfruttate efficacemente in anelli di retroazione per il controllo dell'attuatore SMA.

Malgrado nel corso degli ultimi decenni siano stati messi a punto molti attuatori a memoria di forma [3, 8-11], si osserva ancora la mancanza di formule e di una metodologia utili alla progettazione razionale di attuatori SMA [12]. Inoltre, per via dell'isteresi propria degli attuatori SMA, è ancora aperto il problema del controllo efficace ed economico di attuatori SMA. Il presente lavoro cerca di dare una risposta a queste problematiche fornendo una metodologia razionale di progetto per attuatori a memoria di forma (Paragrafo 2) e proponendo un metodo di controllo efficace ed economico (Paragrafo 3). La descrizione di alcuni prototipi di attuatori a memoria di forma, ispirati alle linee guida proposte, conclude il lavoro (Paragrafo 4).

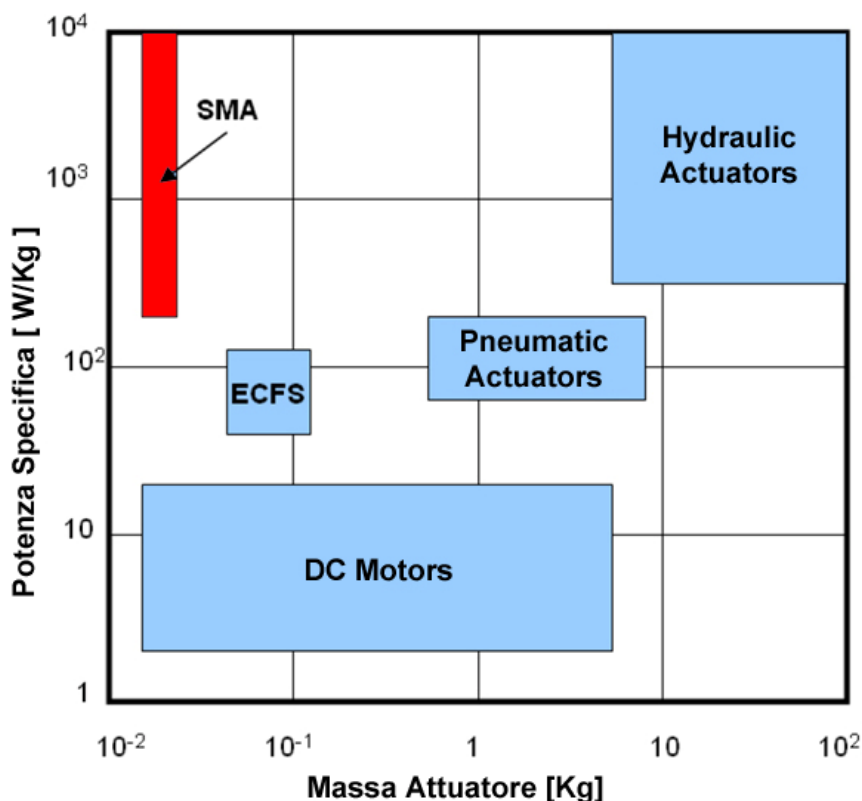


Figura 4. Grafico comparativo delle potenze specifiche di attuatori

2. FORMULE DI DIMENSIONAMENTO

Risulta chiaro che il materiale a memoria di forma, oltre a subire un ciclo termico necessario all'attuazione, necessita di un elemento meccanico in grado di fornire una forza di ripristino [12]. Tale forza consente che lo SMA venga riarmato per poter effettuare un nuovo ciclo di attuazione. Tipicamente, l'elemento attivo viene contrapposto a forze costanti quali gravi, dispositivi pneumatici, molle neg'ator, molle magnetiche o altri meccanismi a forza costante (Fig. 5 (a)), oppure a molle convenzionali (Fig. 5 (b)), o ancora, si utilizza come elemento di ripristino un altro elemento SMA (Fig. 5 (c)).

Nei tre casi caratteristici illustrati nella Fig. 5 si fa ricorso a molle elicoidali, ma è possibile utilizzare qualsiasi altro tipo di elemento elastico. In particolare, i fili a memoria di forma costituiscono il formato più efficiente in termini di potenza specifica e di banda passante [3]. L'inconveniente rilevante legato all'utilizzo di fili SMA è la notevole lunghezza assiale che i fili stessi devono possedere per garantire corse sufficienti significative senza superare deformazioni che pregiudicherebbero la vita a fatica del componente [1,3].

Mediante considerazioni di equilibrio, sono state messe a punto equazioni di progetto utili al dimensionamento degli attuatori SMA nei tre casi tipici di sistema di ripristino: forza costante di ripristino, molla convenzionale di ripristino e molla SMA [12].

Si è assunto che ciascun attuatore debba possedere una corsa utile di progetto Δx e che lungo l'intera corsa l'attuatore debba vincere una forza dissipativa F_F definita come la massima forza che l'attuatore incontra lungo Δx . La forza F_F agisce sempre in direzione opposta al moto dell'attuatore.

Si è considerato un comportamento elastico lineare del materiale, tanto per il comportamento austenitico (ovvero per temperature superiori ad A_f), quanto per il comportamento martensitico (per temperature inferiori ad M_f). Nel normale campo di utilizzo degli SMA come attuatori, la prima assunzione è fedele alla realtà, mentre la seconda non è automaticamente verificata, per via della pseudoplasticità propria della martensite. Dunque, si rende necessaria una spiegazione. In Fig. 6 è presentato il comportamento semplificato tensione – deformazione di un materiale a memoria di forma, nel quale sono riportate le curve di trazione ad alta temperatura (curva H) ed a bassa temperatura (curva L). Un attuatore difficilmente si trova a lavorare oltre il punto C, dunque in campo pseudoelastico. Quindi l'assunzione di linearità in condizione di alta temperatura è rispondente al vero ed il modulo elastico di riferimento utilizzato per sviluppare formule di progetto è E_a . Viceversa, per quanto riguarda il percorso a bassa temperatura, si entra già con basse deformazioni in campo pseudoplastico. Dunque, è necessario fissare un valore di deformazione massima ammissibile, ε_{admS} , che sia la più elevata possibile per sfruttare appieno le proprietà del materiale, garantendo al contempo il rispetto dei vincoli di vita a fatica dell'attuatore. Scelto ε_{admS} , è possibile ricavare il modulo elastico secante E_m . Poiché le formule di progetto mirano al dimensionamento di attuatori SMA, prescindendo dalle posizioni intermedie di funzionamento, è lecito considerare un comportamento lineare sia in condizione di alta temperatura (segmento OC), sia di bassa temperatura (segmento OA').

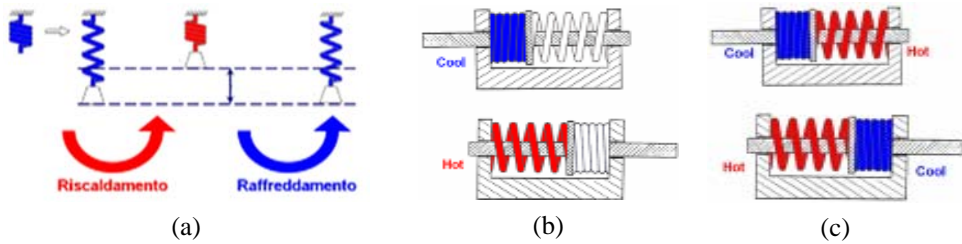


Figura 5. Tre sistemi di ripristino per azionamenti a memoria di forma: (a) forza costante; (b) molla convenzionale; (c) molla a memoria di forma.

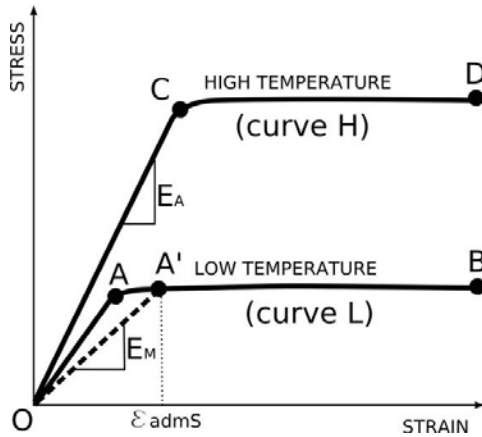


Figura 6. Comportamento semplificato di un materiale a memoria di forma

Vengono definite tre costanti adimensionali, utili a semplificare i calcoli e dotate anche di un significato fisico.

La prima costante, s_1 , è una proprietà della lega a memoria di forma ed è definita come il rapporto tra la rigidezza a caldo, K_{1sh} o K_{2sh} , e la rigidezza a freddo, K_{1sc} o K_{2sc} , rispettivamente delle molla SMA 1 (molla principale) o della molla SMA 2 (molla di ripristino del caso illustrato in Fig. 9, relativo a due SMA contrapposti). Poiché si ipotizza che le caratteristiche delle molle siano indipendenti dalla corsa, s_1 è definibile anche come il rapporto tra il modulo elastico a caldo, E_a o G_a , ed a freddo, E_m o G_m .

$$s_1 = \frac{K_{1SH}}{K_{1SC}} = \frac{K_{2SH}}{K_{2SC}} = \frac{E_A}{E_M} = \frac{G_A}{G_M} \quad (1)$$

La Eqn. 1 implica che, se l'attuatore è costituito da due molle SMA contrapposte, si assume che il materiale col quale sono realizzate sia il medesimo.

Si osserva che s_1 è una figura di merito della lega a memoria di forma, ed in particolare tanto più questo valore è elevato, tanto più marcato sarà l'effetto di memoria di forma.

La seconda costante, s_2 , è utile per i casi di attuatori in cui il ripristino è fornito da una molla (convenzionale o SMA) ed è definita come rapporto tra il valore minimo di rigidezza assunto dalla molla di ripristino (molla 2) e la rigidezza a freddo della molla SMA primaria (molla 1).

$$s_2 = \frac{K_{2min}}{K_{1SC}} \quad (2)$$

Il valore minimo di rigidezza della molla 2, K_{2min} , è pari alla semplice rigidezza della molla, K_{2c} , se questa è una molla convenzionale, mentre è pari alla rigidezza a freddo della molla di ripristino, K_{2sc} , se la molla di ripristino è una molla SMA.

La terza costante introdotta, s_F , serve a quantificare l'influenza della forza dissipativa F_F sull'attuatore ed è definita come il rapporto tra la forza dissipativa, F_F , ed il prodotto della rigidezza a freddo della molla SMA principale, K_{1sc} , e la deflessione a freddo della stessa, $(L_c - L_{01})$, con L_c massima lunghezza assunta dalla molla 1 ed L_{01} lunghezza libera della stessa.

$$s_F = \frac{F_F}{K_{1SC} \cdot (L_c - L_{01})} \quad (3)$$

La procedura di progetto proposta è divisa in due parti: una prima parte nella quale si sono ricavate equazioni aventi validità generale, che prescinde dalla tipologia di molla utilizzata, ed una seconda parte nella quale si sono calcolate formule specifiche, in grado di utilizzare l'output della prima parte per progettare molle reali, una volta scelta la tipologia di molla (es. molle elicoidali, fili, eccetera).

Di seguito si riportano le equazioni finali relative alla parte avente validità generale, per i tre casi caratteristici in esame. Per la procedura utilizzata per ricavare tali formule e per le formule specifiche si rimanda a [12] ed ai consueti testi di costruzione di macchine [13].

Ripristino fornito da una forza costante

L'espressione della corsa dell'attuatore SMA nel quale l'elemento attivo è contrastato da una forza costante di ripristino è la seguente:

$$\Delta x = (L_c - L_{01}) \frac{s_1 - 1 - 2s_F}{s_1} \quad (4)$$

Il valore di s_F deve essere scelto maggiore di 0 e inferiore al seguente valore critico, che determina il bloccaggio dell'attuatore:

$$s_{Fcr} = \frac{s_1 - 1}{2} \quad (5)$$

Infine, la forza di ripristino, F_0 , deve essere scelta secondo la seguente espressione:

$$F_0 = F_F \cdot \frac{s_F + 1}{s_F} \quad (6)$$

La procedura utile alla progettazione dell'attuatore SMA con ripristino fornito da una forza costante è la seguente:

- si assumono come input di progetto la corsa desiderata dell'attuatore, Δx , e la forza dissipative di progetto, F_F ;
- si seleziona un materiale a memoria di forma ed una deformazione massima ammissibile, ε_{admS} o γ_{admS} , dopodiché il parametro $s_1 = E_A/E_M = G_A/G_M$ è noto;
- si sceglie un valore di s_F inferiore al valore critico fornito da Eqn. (5);
- si calcola la forza di ripristino, F_0 , mediante Eqn. (6);
- si calcola la massima deflessione della molla SMA, $L_C - L_{01}$, mediante Eqn. (4);
- si ricava la rigidezza a freddo della molla SMA, K_{Isc} , mediante Eqn. (3).

Questa procedura garantisce che l'attuatore progettato soddisfi le specifiche di input in termini di corsa utile, forza da vincere e vita a fatica, sfruttando un elemento SMA generico, caratterizzato da una rigidezza K_{Isc} e freccia massima, $L_C - L_{01}$, sotto l'azione del carico esterno e di una forza di ripristino, F_0 .

A questo punto è possibile decidere la tipologia di molla da sfruttare per realizzare la molla SMA (elicoidale, fili, Belleville, eccetera) e progettare tale molla nel dettaglio, avente le caratteristiche di rigidezza e freccia massima richieste, secondo [12] e secondo le normali formule della costruzione di macchine [13].

Ripristino fornito da una molla convenzionale

L'espressione della corsa dell'attuatore SMA nel quale l'elemento attivo è contrastato da una molla costante è la seguente:

$$\Delta x = (L_C - L_{01}) \frac{s_1 - 1 - 2s_F}{s_1 + s_2} \quad (7)$$

Il valore di s_F deve essere scelto maggiore di 0 e inferiore al seguente valore critico, che determina il bloccaggio dell'attuatore:

$$s_{Fcr} = \frac{s_1 - 1}{2} \quad (8)$$

La predeformazione, p , da fornire al sistema (molla SMA e molla convenzionale) deve essere scelta secondo la seguente espressione:

$$p = (L_C - L_{01}) \frac{s_2 + s_F + 1}{s_2} \quad (9)$$

La deflessione massima cui è soggetta la molla convenzionale può essere espressa come segue:

$$p - (L_H - L_{01}) = (L_C - L_{01}) \cdot \frac{s_1 + s_1 \cdot s_2 + s_1 \cdot s_F - s_2 \cdot s_F}{s_2 \cdot (s_1 + s_2)} \quad (10)$$

Infine, si può dimostrare come esista un valore di s_2 ottimale che massimizza il rapporto $\Delta x / p$, capace dunque di limitare l'ingombro totale dell'attuatore:

$$s_2^* = \sqrt{s_1 \cdot (s_F + 1)} \quad (11)$$

La procedura utile alla progettazione dell'attuatore SMA con ripristino fornito da una molla convenzionale è la seguente:

- si assumono come input di progetto la corsa desiderata dell'attuatore, Δx , e la forza dissipative di progetto, F_F ;
- si seleziona un materiale a memoria di forma ed una deformazione massima ammissibile, ε_{admS} o γ_{admS} , dopodiché il parametro $s_1 = E_A/E_M = G_A/G_M$ è noto;
- si sceglie un valore di s_F inferiore al valore critico fornito da Eqn. (8);
- si sceglie il valore più piccolo possibile di s_2 (per massimizzare lo sfruttamento della molla SMA) o il valore ottimale fornito da Eqn. (11) (per ridurre l'ingombro dell'attuatore);
- si calcola la massima deflessione della molla SMA, $L_C - L_{01}$, mediante Eqn. (7);
- si ricava la rigidità a freddo della molla SMA, K_{Isc} , mediante Eqn. (3);
- si calcola la predeformazione complessiva, p , mediante Eqn. (9);
- si calcola la rigidità della molla convenzionale, K_{2c} , mediante Eqn. (2);
- si calcola la massima deflessione della molla convenzionale mediante Eqn. (10).

Questa procedura garantisce che l'attuatore progettato soddisfi le specifiche di input in termini di corsa utile, forza da vincere e vita a fatica, sfruttando un elemento SMA generico, caratterizzato da una rigidità K_{Isc} e freccia massima, $L_C - L_{01}$, sotto l'azione del carico esterno e di una molla convenzionale di ripristino avente rigidità K_{2c} e freccia massima $p - (L_H - L_{01})$. Al sistema viene impartita una predeformazione p .

A questo punto è possibile decidere la tipologia di molla da sfruttare per realizzare la molla SMA e la molla convenzionale (elicoidale, fili, Belleville, eccetera) e progettare le molle nel dettaglio, aventi le caratteristiche di rigidità e frecce massime richieste, secondo [12] e secondo le normali formule della costruzione di macchine [13].

Due molle SMA contrapposte

L'espressione della corsa dell'attuatore costituito da due molle SMA contrapposte è la seguente:

$$\Delta x = (L_C - L_{01}) \cdot \frac{(s_1 - 1 - s_F) \cdot (s_1 + 1)}{s_1 \cdot (s_1 + s_2)} \quad (12)$$

Il valore di s_F deve essere scelto maggiore di 0 e inferiore al seguente valore critico, che determina il bloccaggio dell'attuatore:

$$s_{Fcr} = (s_1 - 1) \quad (13)$$

La predeformazione, p , da impartire al sistema (molla SMA e molla convenzionale) deve essere scelta secondo la seguente espressione:

$$p = (L_C - L_{01}) \cdot \frac{s_1 \cdot s_2 + s_F + 1}{s_1 \cdot s_2} \quad (14)$$

La deflessione massima cui è soggetta la molla convenzionale può essere espressa come segue:

$$p - (L_H - L_{01}) = (L_C - L_{01}) \cdot \frac{s_1 \cdot s_2 - s_2 \cdot s_F + s_F + 1}{s_2 \cdot (s_1 + s_2)} \quad (15)$$

Infine, si può dimostrare come esista un valore di s_2 ottimale, che massimizza il rapporto $\Delta x / p$, capace dunque di ridurre l'ingombro totale dell'attuatore:

$$s_2^* = \sqrt{s_F + 1} \quad (16)$$

La procedura utile alla progettazione dell'attuatore SMA con ripristino fornito da una molla convenzionale è la seguente:

- si assumono come input di progetto la corsa desiderata dell'attuatore, Δx , e la forza dissipativa di progetto, F_F ;
- si seleziona un materiale a memoria di forma ed una deformazione massima ammissibile, ε_{admS} o γ_{admS} , dopodiché il parametro $s_1 = E_A/E_M = G_A/G_M$ è noto;
- si sceglie un valore di s_F inferiore al valore critico fornito da Eqn. (13);
- si sceglie il valore più piccolo possibile di s_2 (per massimizzare lo sfruttamento della molla SMA) o il valore ottimale fornito da Eqn. (16) (per ridurre l'ingombro dell'attuatore);
- si calcola la massima deflessione della molla SMA, $L_C - L_{01}$, mediante Eqn. (12);

- si ricava la rigidezza a freddo della molla SMA, K_{1sc} , mediante Eqn. (3);
- si calcola la predeformazione complessiva, p , mediante Eqn. (14);
- si calcola la rigidezza della molla convenzionale, K_{2c} , mediante Eqn. (2);
- si calcola la massima deflessione della molla convenzionale mediante Eqn. (15).

Questa procedura garantisce che l'attuatore progettato soddisfi le specifiche di input in termini di corsa utile, forza da vincere e vita a fatica, sfruttando un elemento SMA generico, caratterizzato da una rigidezza K_{1sc} e freccia massima, $L_C - L_{01}$, sotto l'azione del carico esterno e di una seconda molla SMA di ripristino avente rigidezza a freddo K_{2sc} e freccia massima $p - (L_H - L_{01})$. Al sistema viene impartita una predeformazione p .

A questo punto è possibile decidere la tipologia di molla da sfruttare per realizzare le due molle SMA (elicoidale, fili, Belleville, eccetera) e progettare le molle nel dettaglio, aventi le caratteristiche di rigidezza e frecce massime richieste, secondo [12] e secondo le normali formule della costruzione di macchine [13].

3. CONTROLLO DI POSIZIONE DI ATTUATORI SMA

In ambito industriale, una esigenza che si riscontra di frequente è il controllo di posizione degli attuatori, ossia la possibilità di determinare direttamente con l'attuatore la posizione lineare o angolare dell'intero sistema.

Gli attuatori basati su tecnologia SMA, come detto, annoverano tra i difetti principali le caratteristiche di non linearità e di isteresi, che si traducono in difficoltà di controllo.

Il letteratura sono disponibili diversi lavori [14-17] sul controllo di posizione di attuatori a memoria di forma, che prevedono tutti l'introduzione di un sensore di posizione per la valutazione dell'errore di posizionamento.

Poiché la tecnologia dell'attuazione SMA è utilizzata per lo più nel settore delle microattuazioni, l'introduzione di un sensore di posizione porta molto spesso ad un costo non competitivo del prodotto SMA. Qualora la precisione richiesta nel posizionamento lo consenta, per abbattere il costo è utile evitare di introdurre nel sistema di controllo sensori di posizione, ossia è utile realizzare un controllo del tipo in catena aperta.

La scelta di un controllo in catena aperta impone di individuare tutte le grandezze fisiche che possono interagire con la grandezza che si vuole controllare, in questo caso la posizione. Nel caso di un attuatore SMA, la posizione istantanea dell'attuatore dipende dalla temperatura dell'elemento/i SMA (molle, fili, etc.) e dallo sforzo ad esso applicato. Entrando più nel dettaglio, se l'elemento SMA è attivato sfruttando l'effetto Joule, la sua temperatura dipende dalla tensione che applichiamo ai suoi capi, dalla sua resistenza elettrica e dalla temperatura ambiente in cui l'attuatore è inserito. Come tutte le proprietà fisiche degli SMA, anche la resistività e quindi la resistenza elettrica dipende dalla temperatura e dalla sollecitazione meccanica applicata all'elemento, in modo non lineare e soprattutto isteretico. Dunque, la posizione di un attuatore SMA dipende in modo complesso da molti fattori.

Un caso particolare di funzionamento dell'attuatore SMA si presenta quando il carico meccanico da vincere è costante nel tempo, anche se non necessariamente costante lungo la corsa. Se questa condizione è verificata, è possibile trascurare tutte le dipendenze delle proprietà fisiche delle leghe a memoria di forma dalla sollecitazione meccanica applicata, in quanto questa risulta essere sempre la stessa. Eseguendo una caratterizzazione sperimentale dell'intero sistema (attuatore – sistema movimentato), è possibile rilevare

l'andamento di tutte le grandezze di interesse al variare della posizione assunta dall'attuatore. Ovviamente in questa fase di caratterizzazione si dovrà utilizzare un sensore per rilevare il reale spostamento del sistema. Il risultato finale di questa caratterizzazione è dunque una sorta di mappatura delle grandezze di interesse in funzione della posizione assunta dallo stesso.

Ora, poiché ci si è svincolati dalla sollecitazione meccanica, la posizione dipenderà unicamente dalla temperatura a cui si trova l'elemento a memoria di forma. In condizioni stazionarie, tale elemento deve risultare termicamente in equilibrio: il calore fornito dall'esterno tramite il passaggio di corrente elettrica, deve essere uguale al calore che l'elemento dissipa per convezione con l'aria. Dunque, in condizioni stazionarie si può ricavare che la temperatura dell'elemento SMA è pari a:

$$T_{SMA} = \frac{P_d}{A_{SMA} \cdot h_{SMA}} + T_{\infty} \quad (17)$$

dove A_{SMA} è la superficie di scambio termico, h_{SMA} il coefficiente di convezione, T_{SMA} la temperatura dell'elemento SMA e T_{∞} la temperatura ambiente.

Considerando costante l'area di scambio termico e il coefficiente di convezione, la temperatura dell'elemento dipende quindi esclusivamente dalla potenza fornita per effetto Joule e dalla temperatura ambiente.

Dalla Eqn. 17 emerge che la potenza termica fornita può essere utilizzata efficacemente come parametro per il controllo della temperatura e dunque della posizione, vista la diretta influenza che questa ha sulla temperatura dell'elemento. Poiché in generale la temperatura ambiente non è nota a priori, per eseguire il controllo di posizione sarà necessario monitorarla ed attuare, al variare di questa, opportune azioni correttive.

Eseguendo dunque una caratterizzazione sperimentale dello specifico attuatore in condizioni stazionarie, e sottoposto al carico meccanico di progetto (invariante nel tempo) è dunque possibile ricavare per una data temperatura ambiente, il valore della potenza elettrica da fornire all'elemento per ottenere lo spostamento desiderato, aggirando tutte le dipendenze non lineari delle caratteristiche fisiche dei materiali a memoria di forma.

In Fig. 11 è mostrata un esempio di curva potenza – posizione per un attuatore a tecnologia SMA. Il diagramma è stato ricavato dalla caratterizzazione sperimentale dell'attuatore di Fig. 12. L'attuatore da noi realizzato è composto da un filo SMA di diametro 0.150 mm e lunghezza 50mm, contrapposto ad una molla elicoidale tradizionale. Il tutto è montato direttamente sul sensore di posizione magnetostrittivo Gefran MK4A, che pertanto svolge la duplice funzione di telaio e di sensore di posizione, consentendo di eliminare i giochi di una eventuale catena cinematica.

Il diagramma è stato realizzato aumentando (tratto in color rosso) e diminuendo (tratto in color blu) con velocità di incremento costante ed estremamente bassa (0.15W/min) la potenza applicata al filo SMA e registrando la conseguente posizione dell'attuatore.

L'acquisizione della posizione assunta dall'attuatore ed il controllo della rampa di potenza elettrica sono stata eseguite mediante una scheda di acquisizione dati National Instrument USB 6251 e un software di gestione dedicato sviluppato in ambiente LabView 8.5.

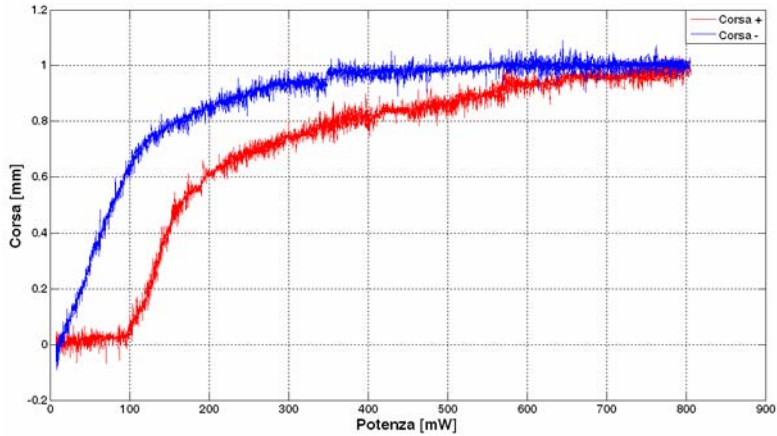


Figura 11. Andamento sperimentale potenza – posizione dell'attuatore di Fig. 12

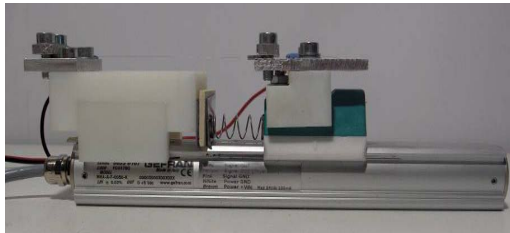


Figura 12. Attuatore SMA controllato in posizione

Il segnale di riferimento della potenza elettrica generato dalla scheda National è applicato al filo tramite una scheda elettronica di potenza e controllo appositamente realizzata (Fig. 13).

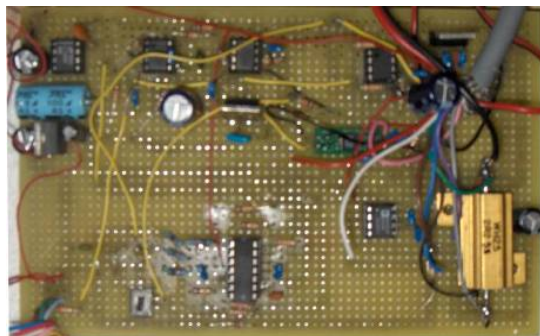


Figura 13. Scheda elettronica di potenza e di controllo

Il prototipo della scheda elettronica, realizzata con componenti discreti, implementa tre funzioni fondamentali per il controllo dell'attuatore:

1. misurare la corrente che attraversa il filo SMA mediante una resistenza di shunt convertirla in un segnale di tensione debitamente amplificato;
2. retroazionare il valore di tensione effettivamente applicato all'elemento SMA, mediante un anello di controllo con azione proporzionale-integrale rispetto alla differenza che questa grandezza ha nei confronti del valore desiderato. Questo valore di riferimento viene stabilito dall'uscita di un secondo anello di retroazione, a monte del precedente, che realizza a sua volta un'azione proporzionale-integrale sull'errore della potenza attuata rispetto alla potenza desiderata applicata all'elemento SMA;
3. calcolare la potenza effettivamente dissipata dall'elemento SMA, andando ad eseguire il prodotto fra i segnali analogici di tensione applicata e di corrente misurata mediante lo shunt. Questa funzione è svolta dall'integrato Analog Device AD633 e rappresenta il riferimento per l'anello di potenza.

In Fig. 14 è visibile lo schema a blocchi complessivo della scheda elettronica di controllo e potenza, in cui si possono notare i due anelli di controllo innestati uno dentro all'altro. L'anello interno si occupa dell'inseguimento della tensione desiderata sul filo e quello esterno della potenza elettrica dissipata.

Il risultato della caratterizzazione sperimentale è la relazione fra la potenza applicata e la posizione assunta dall'attuatore. Come si osserva in Fig. 11, il sistema oltre ad essere non lineare è anche dinamico, ossia la posizione non dipende solo dalla potenza applicata in quell'istante, ma anche dalla potenza applicata precedentemente. Invertendo la relazione, ossia implementando la relazione tra posizione e potenza applicata è possibile risalire alla potenza da applicare all'elemento SMA per raggiungere una posizione assegnata. In altre parole dalla relazione inversa si ricava il segnale di riferimento dell'anello di retroazione della potenza.

L'implementazione di questa relazione sperimentale può avvenire in differenti modi. Un primo approccio è quello di approssimare i due rami di salita e di discesa, con una serie più o meno fine di tratti lineari (approssimazione bilineare o trilineare).

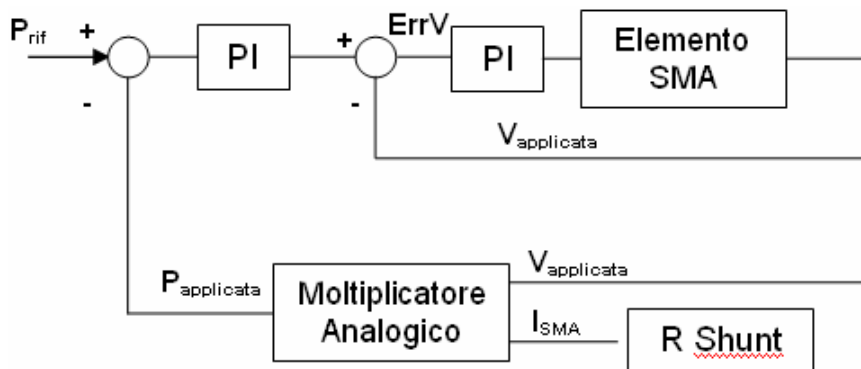


Figura 14. Scheda a blocchi scheda elettronica di funzionamento e controllo

Ricavando i parametri dei vari tratti lineari dall'interpolazione dei tracciati sperimentali è quindi possibile implementare la legge multilineare sia in analogico, mediante reti di amplificatori operazionali, sia in digitale, utilizzando un microcontrollore per il calcolo della potenza da applicare. Gli svantaggi principali del primo approccio sono innanzitutto l'imprecisione intrinseca legata all'approssimazione multilineare oltre alla difficoltà di implementare una correzione della potenza da applicare in funzione della temperatura ambiente. La seconda metodologia consiste invece nell'utilizzare un microcontrollore digitale dotato di memoria interna ed inserire per punti l'intera caratteristica sperimentale individuata. Questa modalità, seppur più costosa nel caso di bassa numerosità di pezzi consente di automatizzare la procedura di caratterizzazione sperimentale dell'attuatore e implementazione della stessa nel controllo oltre che permettere di memorizzare più caratteristiche posizione-potenza al variare della temperatura ambiente. In particolare questo ultimo vantaggio è decisivo per poter realizzare un controllo di posizione preciso e soprattutto robusto al variare delle condizioni ambientali.

Per eseguire la correzione della potenza da attuare in funzione della temperatura ambiente è necessario equipaggiare l'attuatore con un sensore di temperatura posto nella zona dell'elemento SMA, in modo da registrare la temperatura alla quale l'ambiente scambia calore con l'elemento attivo.

Dalla conoscenza di questa grandezza è quindi possibile applicare azioni correttive che vadano a modificare la relazione posizione – potenza. Una prima strada è quella di applicare un coefficiente di correzione calcolato come funzione più o meno complessa della temperatura ambiente. Il vantaggio di questa soluzione è che la dipendenza può essere individuata una volta per tutte su un solo esemplare. Per contro la precisione che si ottiene è ridotta. Il secondo metodo consiste invece nell'eseguire più caratterizzazioni sperimentali dell'attuatore a differenti valori di temperatura ambiente (camera climatica) e realizzare una matrice tridimensionale da inserire nel microcontrollore.

In Fig. 15 è illustrato lo schema a blocchi del controllo di posizione realizzato. Dal valore obiettivo di posizione e dalla temperatura ambiente a cui si trova l'attuatore, il microcontrollore interpola la matrice tridimensionale e ricava il valore della potenza termica da applicare all'elemento a memoria di forma. Tale valore di riferimento viene quindi utilizzato come input dell'anello di retroazione della potenza.

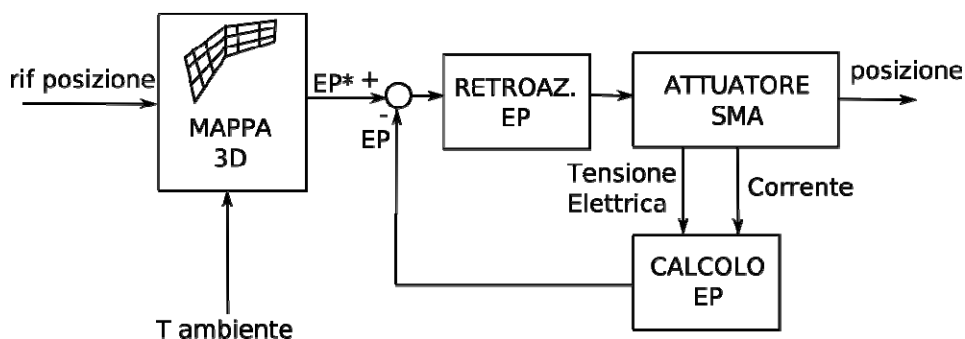


Figura 15. Schema a blocchi delle funzioni del microcontrollore

La matrice tridimensionale risulta composta da due sottomatrici, che rappresentano i valori di potenza da applicare a seconda della potenza e della temperatura ambiente dello stato precedente. Si possono presentare 3 differenti casi:

1. Temperatura ambiente e posizione obiettivo invariate rispetto allo stato precedente: il controllore riapplica la potenza dello stato precedente;
2. temperatura ambiente diminuita rispetto allo stato precedente e/o posizione obiettivo maggiore rispetto alla posizione precedente: il controllore interpola la sottomatrice relativa al ramo di salita;
3. temperatura ambiente aumentata rispetto allo stato precedente e/o posizione obiettivo minore rispetto alla posizione precedente: il controllore interpola la sottomatrice relativa al ramo di discesa.

La determinazione a quale dei tre casi è riconducibile lo stato attuato dell'attuatore è effettuato continuamente dal microcontrollore, applicando il diagramma di flusso di Fig. 16.

Il sistema di controllo messo a punto è stato utilizzato per il controllo di posizione dell'attuatore di Fig. 12. Per esigenze sperimentali, al posto del microcontrollore si è utilizzato direttamente il pacchetto LabView e scheda di acquisizione, lasciando comunque inalterata la metodologia di controllo.

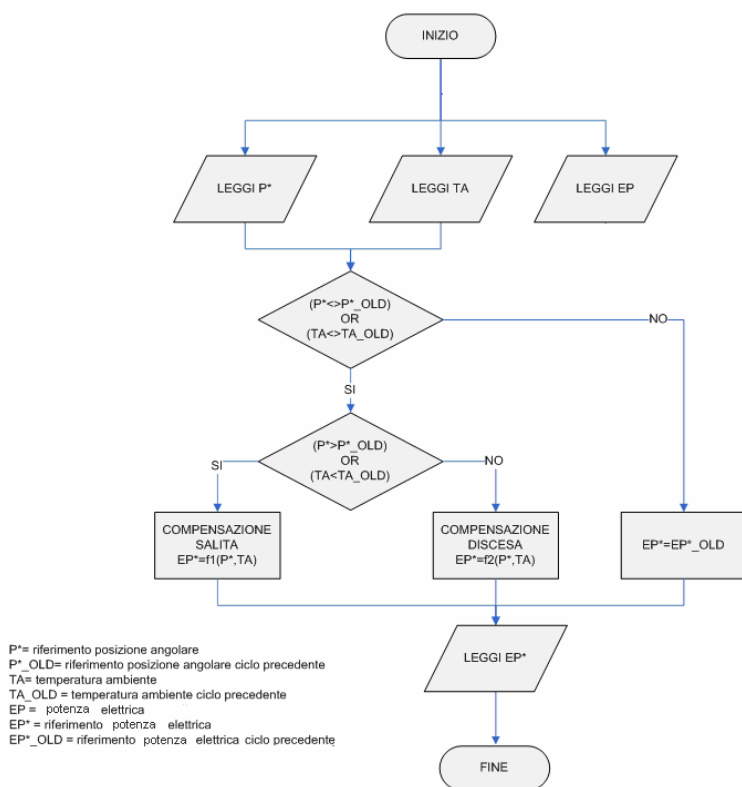


Figura 16. Diagramma di flusso per la determinazione della potenza da attuare

Le prestazioni del sistema di controllo dell'attuatore sono state valutate con due modalità di prova:

1. Valutazione dell'errore di posizionamento in aumento e diminuzione della posizione dell'attuatore
2. Inseguimento da parte dell'attuatore di un segnale di posizione variabile periodicamente

La prima modalità di prova ha permesso di valutare l'errore assoluto di posizionamento (Fig. 17) commesso dal sistema attuatore – controllo durante un movimento monotono dello stesso attuatore (rispettivamente in fase di salita e in fase di discesa). Come si osserva dagli andamenti, l'errore risulta compreso fra ± 0.15 mm, per una corsa totale dell'attuatore di circa 1.5 mm, corrispondente ad un errore $\pm 10\%$. La prestazione è abbastanza soddisfacente per un controllore in catena aperta. La prestazione del controllo è inoltre stata penalizzata da fenomeni di “stick-slip”, attribuibili allo scorrimento del corsoio del sensore magnetostriativo ed evidenti nel tracciato di Fig. 17 nella porzione di corsa 0.8-1.5mm.

La prova di inseguimento di un riferimento ciclico triangolare di posizione ha permesso di valutare l'errore del sistema attuatore-controllore durante le fasi di inversione del moto. Come prevedibile, l'errore risulta più marcato nella fase di inversione, ossia al cambiamento della direzione del moto dell'attuatore, anche se si può notare come durante le fasi monotone si registri un riallineamento fra il segnale di riferimento e la posizione attuata (Fig. 18).

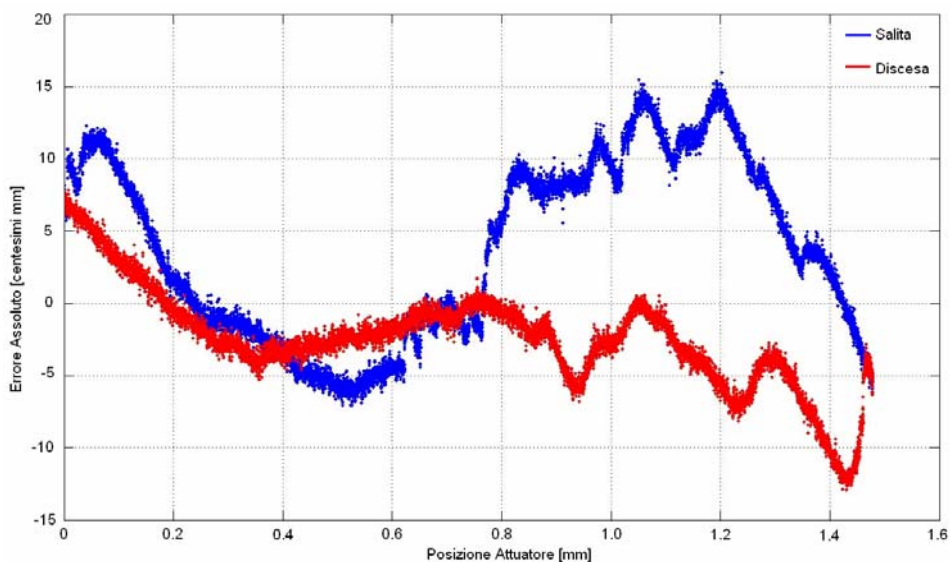


Figura 17. Errore assoluto di posizionamento dell'attuatore in funzione della corsa

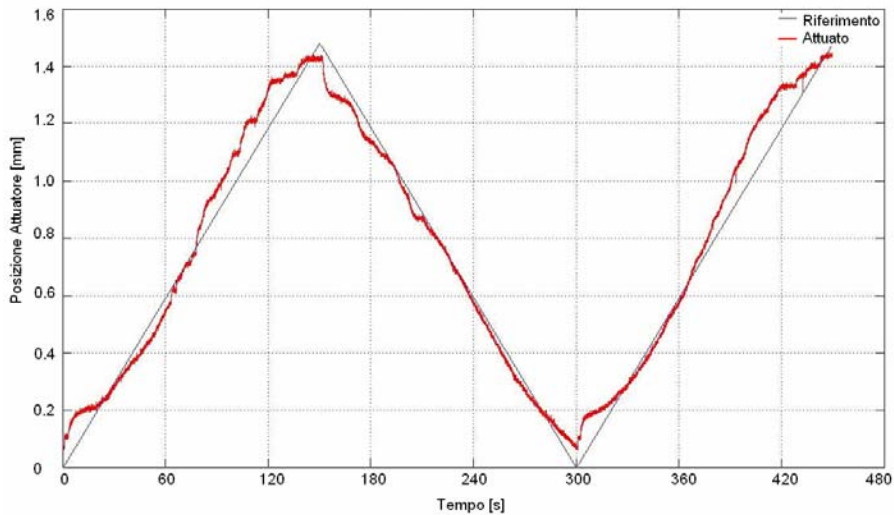


Figura 18. Prestazione dell'attuatore nell'inseguimento di un riferimento triangolare di posizione

4. REALIZZAZIONE DI PROTOTIPI FUNZIONALI A MEMORIA DI FORMA

Oltre al prototipo di attuatore utilizzato per lo sviluppo del controllo di posizione in catena aperta, presso il gruppo di ricerca di costruzione di macchine del DISMI "Dipartimento di Scienze e Metodi dell'Ingegneria" sono stati sviluppati altri dispositivi attuati mediante tecnologia SMA, ed in particolare si è realizzata una valvola oleidraulica SMA per il controllo della portata.

Valvola oleidraulica di controllo della portata

Una interessante applicazione sviluppata, originata da una richiesta industriale, è stata l'attuazione mediante tecnologia a memoria di forma di una valvola oleidraulica per il controllo della portata in impianti oleodinamici. L'obiettivo della ricerca è stato quello di sostituire la normale movimentazione a bobina-solenioide con una attuazione SMA, cercando di mantenere invariate le prestazioni della valvola in termini di tempi di risposta e di capacità di apertura/chiusura e riducendo possibilmente gli ingombri.

La valvola oleidraulica su cui si è intervenuti è una normale valvola servoidraulica ON/OFF, normalmente chiusa, per la regolazione del flusso. La pressione massima di esercizio, nella versione con bobina solenoide è di 310bar. Il principio di funzionamento su cui si basa la valvola è molto semplice. Con riferimento alla Fig. 19a, uno spillo ferromagnetico 1, inserito nel corpo valvola 2 è contrastato da una molla tradizionale 3. Il solenoide 4 esterno al corpo valvola, qualora attivato, genera un campo magnetico capace di attrarre lo spillo e conseguentemente aprire la luce di passaggio dell'olio. In realtà, essendo una servovalvola, lo spillo non agisce direttamente sulla luce primaria, ma ha lo scopo di generare uno squilibrio di pressioni all'interno del corpo valvola che pilota l'apertura del flusso nel condotto primario. Alla cessazione del campo magnetico, la molla

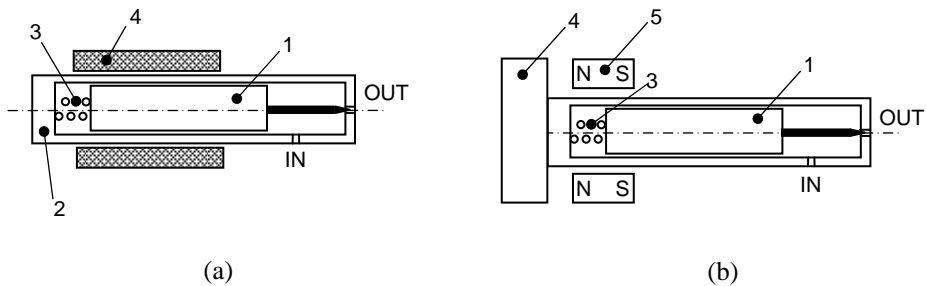


Figura 19. Schema funzionale della valvola servoidraulica di controllo della portata: (a) versione tradizionale attuata mediante bobina-solenioide; (b) versione attuata mediante tecnologia SMA

di contrasto provvede a riportare in posizione lo spillo. A seguito della chiusura della luce lo squilibrio di pressioni cambia di direzione e provoca l'interruzione del flusso.

La conversione dell'attuazione dal sistema bobina-solenioide alla tecnologia SMA è stata particolarmente critica a causa dell'impossibilità di inserire l'elemento a memoria di forma direttamente all'interno del corpo valvola. Ciò è dovuto alla limitata temperatura di funzionamento delle leghe a memoria di forma. Infatti, per poter passare dallo stato austenitico (alta temperatura, attivo) allo stato martensitico (bassa temperatura, non attivo), il materiale deve raffreddarsi sotto M_f . Se la temperatura di lavoro risulta maggiore della temperatura M_f di fine trasformazione martensitica il materiale non viene completante disattivato e quindi il dispositivo non funziona. Inoltre, anche qualora questa condizione sia verificata, il tempo di disattivazione del dispositivo è tanto maggiore quanto più la temperatura ambiente alla quale si trova lo SMA è elevata. Per garantire dinamiche di disattivazione rapide si è quindi evitato che l'elemento a memoria di forma fosse a contatto con il fluido di lavoro, collocando lo SMA all'esterno del corpo valvola. Questa soluzione ha reso necessario trasmettere il moto generato all'esterno del corpo valvola allo spillo otturatore interno. La trasmissione ideata (Fig. 19b) prevede un accoppiamento magnetico fra un magnete esterno 5 coassiale al corpo valvola 2 e lo spillo ferromagnetico interno 1, attraverso un elemento 4 anch'esso ferromagnetico che ha lo scopo di realizzare il circuito magnetico. Il sistema SMA agisce direttamente sul magnete esterno, spostandolo assialmente verso l'elemento 4 e viceversa grazie ad una molla di contrasto. Il movimento del magnete provoca una variazione di riluttanza del circuito magnetico che porta l'otturatore a vincere la molla di contrasto interna e la pressione dell'olio aprendo la luce di passaggio dell'olio. Per ottimizzare le prestazioni del circuito magnetico l'intero corpo valvola è stato ridisegnato ed ottimizzato con l'ausilio di strumenti agli elementi finiti per la simulazione del comportamento magnetico dell'intero circuito (Software FEMM). Oltre alla geometria è inoltre stato modificato il materiale di realizzazione del corpo valvola, passando da un acciaio ad alta resistenza ad un acciaio amagnetico.

Riguardo all'attuazione SMA, sempre per motivi di rapidità di attivazione e disattivazione, si è optato per l'utilizzo di una serie di fili a memoria di forma del diametro di 0.150mm. Dalla corsa richiesta al magnete movimentato si è determinato la lunghezza libera dei fili, fissando la deformazione massima ammissibile nell'ottica di resistere per

almeno 50.000 cicli. Il riarmo dei fili SMA e quindi l'allontanamento del magnete dall'elemento 4 è affidato ad una molla elicoidale tradizionale.

In Fig. 20 si presenta il diagramma di progetto dell'attuazione SMA con l'andamento delle forze agenti sul magnete. Per il calcolo del numero dei fili SMA necessari per l'attuazione si è dovuto considerare oltre alla forza della molla anche la forza magnetica di attrazione del magnete (elemento in rosso in Fig. 21a) con il tappo ferromagnetico realizzato in due mezzi del corpo valvola (in blu e in verde in Fig. 21a), al variare della posizione del magnete

La condizione di funzionamento del dispositivo è garantita dal fatto che la somma algebrica delle forze magnetiche e della forza dei fili SMA sia sempre superiore, in ogni punto della corsa del magnete, alla forza di contrasto nel caso dei fili attivati, e sempre inferiore nel caso dei fili SMA disattivati (Fig. 20).

Il precarico dei fili SMA è regolato ruotando l'esagono del corpo valvola rispetto al telaio del sistema SMA. In questo modo la molla di contrasto viene invece decompressa. Per poter regolare in modo indipendente il precarico dei fili SMA e della molla di contrasto, sono state predisposte tre viti mordenti poste all'estremità inferiore della valvola che, agendo su un piattello (in rosa in Fig. 20), consentono di regolare il precarico della sola molla di contrasto.

La versione definitiva prevede l'utilizzo di 3 fili Nitinol (Af 90°C) di diametro 0.150mm. Avvolgendo i tre fili su un ancoraggio solidale all'elemento mobile portamagnete, si ricavano 6 rami attivi. Il serraggio dei fili SMA è effettuato tramite

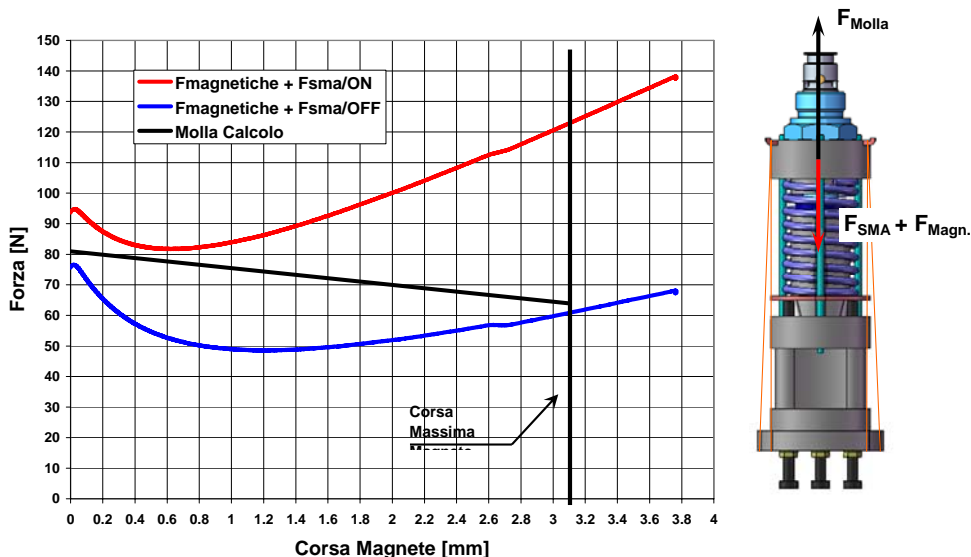


Figura 20. Prestazione dell'attuatore nell'inseguimento di un riferimento triangolare di posizione

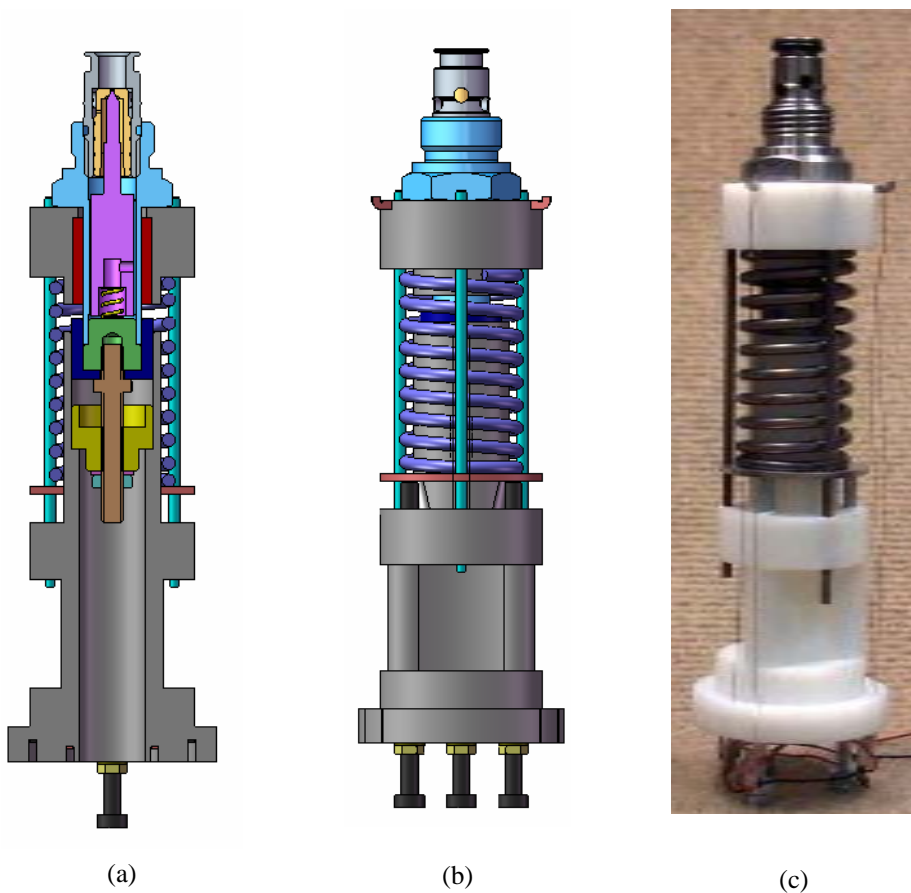


Figura 21 Valvola oleoidraulica attuata mediante tecnologia SMA: (a) vista in sezione del modello CAD 3D; (b) vista frontale del modello CAD 3D; (c) Prototipo fisico della valvola oleoidraulica

crimpatura dei terminali. I crimp sono alloggiati in apposte sedi ricavate all'interno della base (Fig. 23). In queste sedi sono presenti i contatti elettrici per l'alimentazione in parallelo dei fili. Le due piastrine dalla forma a "ragno" sono disposte su due livelli differenti per garantire l'isolamento reciproco. Ogni filo SMA avrà pertanto un terminale sulla piastrina di colore rosso e uno su quella di colore blu (Fig. 23).

A livello elettrico i fili SMA sono alimentati tramite due piastrine di metallo poste all'estremità inferiore della valvola.

Il prototipo definitivo della valvola (Fig. 21c), ed in particolare tutti gli elementi necessari alla trasmissione SMA sono stati realizzati mediante prototipazione rapida con resina ABS.

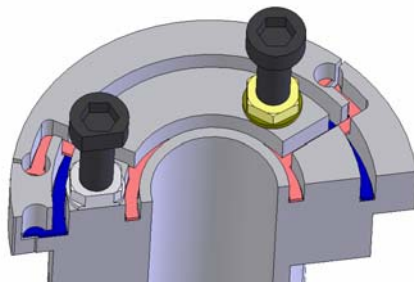


Figura 23. Sedi per alloggiamento dei crimp e relativo contatto elettrico

Il corpo valvola amagnetico è stato invece ricavato per lavorazione tradizionale e a questo si è provveduto a saldare la parte interna del tappo in due parti (Elemento in verde in Fig. 21a) in modo da garantire la tenuta idraulica e la sicurezza del sistema. La parte esterna del tappo invece è solidale alla parte per mezzo di un elemento filettato (marrone in Fig. 21a).

Il prototipo realizzato è stato provato al banco, facendo registrare una pressione di apertura chiusura identica a quella della valvola tradizionale. I tempi di attuazione sono risultati in linea, mentre i tempi di disattivazione sono leggermente aumentati.

L'assorbimento di corrente per l'apertura della valvola è di circa 10W, mentre nella versione tradizionale l'assorbimento nominale è 18W. Per entrambe le soluzioni, la corrente assorbita in mantenimento di apertura è una frazione della corrente massima di apertura, in percentuale maggiore per il sistema SMA.

In Fig. 24 è mostrata la valvola tradizionale equivalente attuata mediante sistema bobina-solenoidale. Come si può notare dal confronto con le Fig. 20-21, l'ingombro della soluzione SMA risulta decisamente più contenuto in direzione radiale. In particolare la soluzione di partenza, registra un ingombro radiale dato dalla dimensione massima della bobina di circa 70mm, mentre la soluzione SMA ha un ingombro radiale uniforme di 48mm. Al contrario in direzione assiale l'aumento dell'ingombro è considerevole. In particolare la soluzione tradizionale ha una lunghezza di 86mm, mentre la versione SMA ha una lunghezza circa doppia pari a 163mm.

L'incremento di dimensione assiale è da imputare principalmente alla necessità di avere una lunghezza dei fili tale rispetto alla quale la corsa richiesta non determini una deformazione maggiore superiore al 4%, valore oltre il quale non sarebbe garantita la durata di 50.000 cicli richiesta al dispositivo.

5. CONCLUSIONI

Sono state sviluppate formule utili al dimensionamento meccanico di attuatori a memoria di forma, distinguendo tre casi caratteristici con i quali viene fornito all'elemento attivo la forza di riarmo: forza costante, molla di contrasto convenzionale e molla di contrasto SMA. La metodologia di progetto delineata conduce al proporzionamento ottimale dell'attuatore attraverso una procedura univoca passo-passo.

Viene proposto un sistema di controllo in catena aperta della posizione dell'attuatore, basato su un doppio anello di retroazione della tensione applicata e della potenza dissipata dall'elemento SMA. La posizione desiderata viene tradotta in una potenza elettrica da

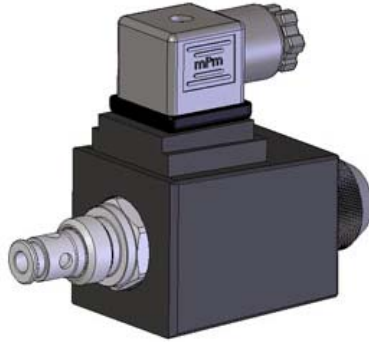


Figura 24. Valvola tradizionale equivalente

fornire, espressa in funzione anche della temperatura ambiente a cui si trova l'attuatore. La relazione posizione-potenza è ricavata sperimentalmente ed inserita all'interno del microcontrollore che implementa il controllo.

Viene inoltre presentata una concreta applicazione industriale relativa alla riprogettazione di una valvola oleoidraulica di controllo della portata nella quale l'attuazione originaria con bobina solenoide viene sostituita da un fascio di fili a memoria di forma.

BIBLIOGRAFIA

- [1] Mavroidis D., Pfeiffer C., Mosley M., and Rutgers U., 2000. *Conventional motors, Shape memory alloys and ElectroRheological fluids*, 1st ed., Vol. 4 of *The Topics on NDE (TONE) Series*. ASNT. Columbus, Ohio, Chap. 5, pp. 189-214.
- [2] Liang, C., and Rogers, C.A., 1997. "One-dimensional thermomechanical constitutive relations for shape memory materials". *Journal of Intelligent Material Systems and Structures*, **8**, pp. 285–302.
- [3] Reynaerts, D., and Van Brussel, H., 1998. "Design aspects of shape memory actuators". *Mechatronics*, **8**, pp. 635-656.
- [4] Mani, G., Feldman, M.D., Patel, D., and Agrawal, C.M., 2007. "Coronary stents: A materials perspective". *Biomaterials*, **28**, pp. 1689-1710.
- [5] Eliades, T., 2007. "Orthodontic materials research and applications: Part 2. Current status and projected future developments in materials and biocompatibility". *American Journal of Orthodontics and Dentofacial Orthopedics*, **131**, pp. 229-237.
- [6] Ma, H., and Cho, C., 2008. "Feasibility study on superelastic SMA damper with re-centring capability". *Materials Science and Engineering A*, **473**, pp. 290-296.
- [7] Choi, S.B., and Hwang, J.H., 2000. "Structural vibration control using shape memory actuators". *Journal of Sound and Vibration*, **231**(4), pp. 1168-1174.
- [8] Kuribayashi, K., 1989. "Millimeter-sized Joint Actuator using a Shape Memory Alloy". *Sensor and Actuators*, **20**, pp. 57-64.

- [9] Khidir, E.A., Mohamed, N.A., Nor, M.J.M. and Mustafa, M.M., 2007. "A new concept of a linear smart actuator". *Sensor and Actuators A*, **135**, pp. 244-249.
- [10] Kim, B., Lee, M.G., Lee, Y.P., Kim, Y. and Lee, G., 2006. "An earthworm-like micro robot using shape memory alloy actuator". *Sensor and Actuators A*, **125**, pp. 429-437.
- [11] Haga, Y., Makishi, W., Iwami, K., Totsu, K., Nakamura, K. and Esashi, M., 2005. "Braille display using SMA coil actuator and magnetic latch". *Sensor and Actuators A*, **119**, pp. 316-322.
- [12] Spinella, I., Dragoni, E., 2008. "Design equations for binary shape memory actuators under dissipative forces". *Journal of Mechanical Engineering Science*, in press.
- [13] Shigley, J.E., Mischke, C.R., and Brown, 2004. *Standard Handbook of Machine Design*, 3rd ed., McGraw-Hill, Columbus, Ohio, Chap. 6.
- [14] C. Dickinson, J. Wen, C., 1998. "Feedback control using shape memory alloy actuators". *Journal of intelligent material systems and structures*, **Vol. 9**, pp. 242-250.
- [15] E. Shameli, A. Alasty, H. Salaarieh, 2005. "Feedback Stability analysis and nonlinear control of a miniature shape memory alloy actuator for precise applications". *Mechatronics*, **Vol. 15**, pp. 471-486.
- [16] Hyo Jik Lee, Jung Ju Lee and Dong Soo Kwon, 2005. "Position Control of an SMA Actuator Using Sliding Mode Control with Time Delay Estimation"
- [17] Yee Harn The, 2003. "A Control System For Achieving Rapid Controlled Motions From Shape Memory Alloy (SMA) Actuator Wires, *The Australian National University, Thesis*

UNA DIMOSTRAZIONE DIDATTICAMENTE EFFICACE DEL TEOREMA DI EULERO-SAVARY

Carlo Innocenti

Dipartimento di Ingegneria Meccanica e Civile,

Università di Modena e Reggio Emilia

E-mail: carlo.innocenti@unimore.it

***Sommario.** Il lavoro propone una dimostrazione alternativa del teorema di Eulero-Savary per il caso di moto piano. La dimostrazione presentata, pur non essendo più sintetica od elegante delle dimostrazioni note, può apparire più convincente di queste agli occhi di chi non ha ancora acquisito una solida padronanza dei concetti fondamentali di cinematica piana.*

***Parole chiave:** profili coniugati, polari del moto, curvatura, involuppo.*

1. INTRODUZIONE

In cinematica piana il teorema di Eulero-Savary stabilisce un legame tra il raggio di curvatura di una curva mobile, i raggi di curvatura delle polari del moto sottese alla traiettoria di corpo rigido della suddetta curva mobile, ed il raggio di curvatura della curva involupata dalla curva mobile. Come caso particolare, quando la curva mobile si riduce ad un punto, il teorema di Eulero-Savary pone in relazione i raggi di curvatura delle polari del moto col raggio di curvatura della traiettoria del punto considerato.

Noto da circa due secoli ([1-2]), il teorema di Eulero-Savary è suscettibile di estensione al caso di moto sferico ([3-4]). In Meccanica applicata alle macchine trova applicazione nello studio del contatto tra profili coniugati (meccanismi a camma, ruote dentate) e nella sintesi dimensionale di sistemi articolati.

Del teorema di Eulero-Savary esistono sia dimostrazioni interamente analitiche [4], sia dimostrazioni analitico-geometriche ([1-3, 5-7]). Essendo per loro natura più astratte, le prime sono meno idonee delle seconde ad un impiego didattico. Pure le seconde possono tuttavia risultare oscure a chi, disponendo unicamente dei più elementari concetti di cinematica, vorrebbe (erroneamente) attribuire al centro di istantanea rotazione di una figura piana in moto rispetto ad un osservatore l'ulteriore ruolo di centro di curvatura delle traiettorie descritte dai punti della figura piana mobile rispetto al suddetto osservatore.

La dimostrazione alternativa del teorema di Eulero-Savary proposta nella presente memoria si avvale di un espediente per guidare il ragionamento logico dello studente: la curva involupante è immaginata solidale alla biella di un quadrilatero articolato piano di dimensioni opportune. La formula di Eulero-Savary emerge quindi dall'individuazione dei requisiti geometrici che un siffatto quadrilatero articolato deve possedere.

Pur avvalendosi di figure esplicative, la dimostrazione presentata è di tipo analitico. Le convenzioni adottate per attribuire un segno algebrico alle diverse grandezze introdotte nel corso della dimostrazione costituiscono la naturale chiave di lettura della formula di Eulero-Savary, nella forma qui proposta.

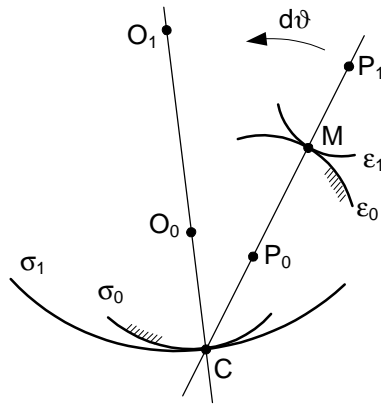


Figura 1. Polari del moto e profili coniugati.

2. IL QUADRILATERO ARTICOLATO AUSILIARIO

La Fig. 1 mostra una curva ε_1 tracciata su un piano mobile π_1 che descrive, rispetto ad un piano fisso π_0 , la traiettoria di corpo rigido definita dal rotolamento della polare mobile σ_1 sulla polare fissa σ_0 . La generica posizione della curva ε_1 rispetto al piano fisso π_0 sia parametrizzata dall'angolo ϑ esprimente l'orientamento di π_1 rispetto a π_0 . Sia ε_0 la curva involupata dalla curva ε_1 durante il moto di quest'ultima rispetto a π_0 . Sia inoltre M il punto di contatto tra ε_0 ed ε_1 . Supponendo note le polari del moto e la curva ε_1 rispetto a π_1 , si voglia determinare – per un generico valore $\bar{\vartheta}$ di ϑ – il centro di curvatura P_0 della curva ε_0 nel punto M .

Per ogni valore del parametro di configurazione ϑ la curva ε_1 ha posizione nota rispetto al piano fisso. Il punto di contatto M tra curva involupante ε_1 e curva involupata ε_0 giace sulla normale ad ε_1 passante per il centro di istantanea rotazione C , punto di contatto tra σ_0 e σ_1 . Su tale retta giace pure, in posizione nota, il centro di curvatura P_1 di ε_1 in M e, in posizione per ora incognita, il centro di curvatura P_0 di ε_0 in M .

Si considerino ora i punti O_0, O_1, P_0 e P_1 , centri di curvatura delle polari σ_0 e σ_1 in C e delle curve ε_0 ed ε_1 in M . Le posizioni dei punti O_0 e P_0 rispetto al piano fisso, come pure le posizioni dei punti O_1 e P_1 rispetto al piano mobile, dipendono, in generale, dal parametro ϑ . Sono quindi generalmente non nulle le derivate prime rispetto a ϑ dei vettori posizione di O_0 e P_0 nel piano fisso e le derivate prime – sempre rispetto a ϑ – dei vettori posizione di O_1 e P_1 nel piano mobile. Per analoghi motivi è generalmente non nulla la derivata prima rispetto a ϑ della distanza tra i punti O_0 ed O_1 e tra i punti P_0 e P_1 .

In corrispondenza all'assegnato valore $\bar{\vartheta}$ del parametro di configurazione ϑ , siano O_0^* e O_1^* i centri di curvatura di σ_0 e σ_1 in C . Siano inoltre P_0^* e P_1^* i centri di curvatura di ε_0 ed ε_1 in M . Essendo relativi ad un ben preciso valore di ϑ , i punti O_0^* e P_0^* sono immobili rispetto al piano fisso, mentre i punti O_1^* e P_1^* sono immobili rispetto al piano mobile

(l'asterisco usato come apice sta qui ad indicare punti che, durante il moto di π_1 rispetto a π_0 , rimangono immobili rispetto a π_0 o π_1).

Come noto, un cerchio osculatore approssima, nell'intorno del punto di osculazione, la curva osculata con un'accuratezza del second'ordine. Ciò implica che lo scostamento tra polare fissa σ_0 e cerchio osculatore a σ_0 in C sia una quantità infinitesima del terz'ordine in $d\vartheta = \vartheta - \bar{\vartheta}$.

Se le polari del moto σ_0 e σ_1 fossero rimpiazzate, per ϑ prossimo a $\bar{\vartheta}$, dai rispettivi cerchi osculatori in C , al variare di ϑ la distanza $\overline{O_0^*O_1^*}$ tra i centri di curvatura O_0^* e O_1^* rimarrebbe rigorosamente costante. Con polari del moto non approssimate da archi di cerchi osculatori, la quantità $\overline{O_0^*O_1^*}$ non è rigorosamente indipendente da ϑ : nell'intorno di $\bar{\vartheta}$ essa è pari alla somma di una quantità costante e di una quantità infinitesima del terz'ordine in $d\vartheta$. Valgono dunque le relazioni seguenti:

$$\frac{d\overline{O_0^*O_1^*}}{d\vartheta} = \frac{d^2\overline{O_0^*O_1^*}}{d\vartheta^2} = 0; \quad \frac{d^3\overline{O_0^*O_1^*}}{d\vartheta^3} \neq 0 \quad (1)$$

Analoghe considerazioni possono essere riproposte con riferimento alle curve ε_0 ed ε_1 ed ai rispettivi centri di curvatura P_0^* e P_1^* nel punto di tangenza M relativo all'assegnato valore $\bar{\vartheta}$ del parametro di configurazione ϑ :

$$\frac{d\overline{P_0^*P_1^*}}{d\vartheta} = \frac{d^2\overline{P_0^*P_1^*}}{d\vartheta^2} = 0; \quad \frac{d^3\overline{P_0^*P_1^*}}{d\vartheta^3} \neq 0 \quad (2)$$

La dimostrazione del teorema di Eulero-Savary esposta nella presente memoria poggia sulla seguente congettura: considerato che il teorema di Eulero-Savary ha per oggetto proprietà del secondo ordine di curve, la variabilità del terz'ordine delle distanze $\overline{O_0^*O_1^*}$ e $\overline{P_0^*P_1^*}$ dovrebbe non essere rilevante ai fini della dimostrazione del teorema.

Il supporre costanti le distanze $\overline{O_0^*O_1^*}$ e $\overline{P_0^*P_1^*}$ nell'intorno di $\bar{\vartheta}$ equivale a collegare i piani π_0 e π_1 con due membri binari rigidi, incernierati in O_0^* e P_0^* al piano π_0 ed in O_1^* e P_1^* al piano π_1 . Viene quindi a costituirsi il quadrilatero articolato $O_0^*O_1^*P_1^*P_0^*$ – qui denominato quadrilatero articolato ausiliario – che ha per telaio il piano fisso π_0 e per biella il piano mobile π_1 (v. Fig. 2).

Il quadrilatero articolato ausiliario non è al momento interamente noto in quanto, dovendo ancora determinare il centro di curvatura in M della curva involupata ε_0 , la posizione di P_0^* è tuttora incognita. Tuttavia alla traiettoria di corpo rigido della biella $O_1^*P_1^*$ del quadrilatero articolato ausiliario è richiesto di approssimare al secondo ordine la traiettoria di π_1 rispetto a π_0 . Sarà appunto tale requisito a consentire di determinare, nella prossima sezione, la posizione del punto P_0^* .

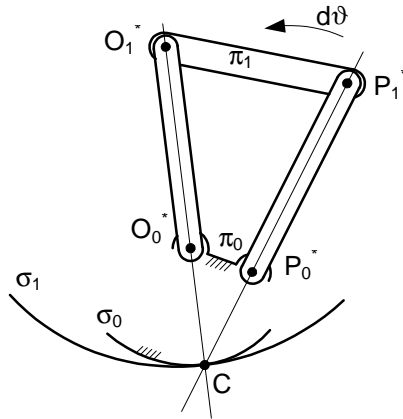


Figura 2. Il quadrilatero articolato ausiliario.

L'ammissibilità del quadrilatero articolato ausiliario quale espediente per la dimostrazione del teorema di Eulero-Savary sarà discussa e provata nella quarta sezione della presente memoria.

3. LA DIMOSTRAZIONE

I bilancieri $O_0^*O_1^*$ e $P_0^*P_1^*$ del quadrilatero articolato ausiliario $O_0^*O_1^*P_1^*P_0^*$ di Fig. 2 guidano la biella $O_1^*P_1^*$ lungo una traiettoria di corpo rigido che – almeno nell'intorno della configurazione individuata dall'angolo $\bar{\vartheta}$ – è assimilabile a quella definita dal rotolamento della polare mobile σ_1 sulla polare fissa σ_0 . Nell'ambito di un'approssimazione del secondo ordine, i due bilancieri del quadrilatero articolato ausiliario costituiscono dunque un'alternativa al vincolo di puro rotolamento tra le polari del moto σ_0 e σ_1 . Ovviamente ciò è vero a patto che il bilanciante $P_0^*P_1^*$ abbia l'estremo P_0^* nel centro di curvatura in M della curva involupata ϵ_0 (v. Fig. 1).

Per determinare la posizione del punto P_0^* sulla retta passante per i punti P_1^* e C occorre imporre che, in riferimento al piano fisso π_0 , la traiettoria di corpo rigido della biella del quadrilatero articolato ausiliario sia identica – localmente ed al second'ordine – alla traiettoria di corpo rigido del piano mobile π_1 rispetto al piano fisso π_0 .

Il centro d'istantanea rotazione C di π_1 rispetto a π_0 coincide già col centro d'istantanea rotazione tra la biella ed il telaio del quadrilatero articolato ausiliario: C è allineato sia con O_0^* ed O_1^* , sia con P_0^* e P_1^* . Sono pertanto simultaneamente soddisfatte le seguenti due condizioni scalari

$$\begin{cases} (O_0^* - C) \times (O_1^* - O_0^*) \cdot \mathbf{k} = 0 \\ (P_0^* - C) \times (P_1^* - P_0^*) \cdot \mathbf{k} = 0 \end{cases} \quad (3)$$

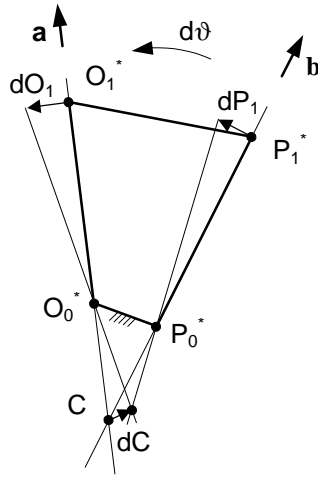


Figura 3. Lo spostamento del centro d'istantanea rotazione tra biella e telaio.

ove \mathbf{k} è il versore ortogonale al piano di figura, diretto verso il lettore.

Assegnati O_0^* , O_1^* , P_1^* e C , le Eq. (3) non definiscono ancora univocamente il punto P_0^* , prescrivendogli semplicemente di giacere sulla retta passante per C e P_1^* . Per determinare P_0^* occorre imporre che, per ϑ appartenente ad un intorno infinitesimo di $\bar{\vartheta}$, il centro d'istantanea rotazione di π_1 rispetto a π_0 continui a cadere sulle polari del moto, nel loro nuovo punto di tangenza.

La Fig. 3 mostra schematicamente i vertici del quadrilatero articolato ausiliario nella configurazione di riferimento ($\vartheta = \bar{\vartheta}$) ed in una configurazione da questa raggiungibile mediante una rotazione infinitesima $d\vartheta$ della biella rispetto al telaio. Nel piano π_0 i vertici O_0^* e P_0^* restano immobili, mentre i vertici O_1^* e P_1^* (immobili nel piano π_1) subiscono rispetto a π_0 gli spostamenti infinitesimi dO_1 e, rispettivamente, dP_1 .

Considerando che la biella del quadrilatero articolato ausiliario subisce, rispetto a π_0 , la rotazione infinitesima $d\vartheta$ attorno al punto C , è facile determinare gli spostamenti infinitesimi dei suoi punti O_1^* e P_1^* :

$$\begin{aligned}
 dO_1 &= d\vartheta \mathbf{k} \times (O_1^* - C) \\
 dP_1 &= d\vartheta \mathbf{k} \times (P_1^* - C)
 \end{aligned}
 \tag{4}$$

Se \mathbf{a} e \mathbf{b} sono due versori paralleli ai lati $O_0^*O_1^*$ e $P_0^*P_1^*$ del quadrilatero articolato ausiliario, con verso scelto ad arbitrio, si può porre

$$\begin{aligned} (O_j^* - C) &= r_j \mathbf{a} & (j = 0, 1) \\ (P_j^* - C) &= s_j \mathbf{b} & (j = 0, 1) \end{aligned} \quad (5)$$

da cui incidentalmente consegue

$$\begin{aligned} r_j &= (O_j^* - C) \cdot \mathbf{a} & (j = 0, 1) \\ s_j &= (P_j^* - C) \cdot \mathbf{b} & (j = 0, 1) \end{aligned} \quad (6)$$

Grazie alle Eq. (5), le Eq. (4) possono essere così riscritte

$$\begin{aligned} dO_1 &= d\vartheta r_1 \mathbf{k} \times \mathbf{a} \\ dP_1 &= d\vartheta s_1 \mathbf{k} \times \mathbf{b} \end{aligned} \quad (7)$$

Per determinare lo spostamento dC rispetto al piano π_0 del centro di istantanea rotazione tra la biella ed il telaio del quadrilatero articolato ausiliario, si differenziano ora le Eq. (3) tenendo presente che O_0^* e P_0^* sono punti immobili rispetto al telaio

$$\begin{cases} \left[-dC \times (O_1^* - O_0^*) + (O_0^* - C) \times dO_1 \right] \cdot \mathbf{k} = 0 \\ \left[-dC \times (P_1^* - P_0^*) + (P_0^* - C) \times dP_1 \right] \cdot \mathbf{k} = 0 \end{cases} \quad (8)$$

Grazie alle Eqq. (4) e (5), le Eq. (8) possono essere poste nella forma seguente

$$\begin{cases} (r_0 - r_1) \mathbf{k} \times \mathbf{a} \cdot dC = r_0 r_1 d\vartheta \\ (s_0 - s_1) \mathbf{k} \times \mathbf{b} \cdot dC = s_0 s_1 d\vartheta \end{cases} \quad (9)$$

Queste due equazioni scalari potrebbero consentire di determinare le due componenti di dC se tutte le altre quantità fossero note. Nel caso qui prospettato non è tuttavia noto s_0 in quanto (v. la seconda delle Eq. (6) per $j=0$) il punto P_0^* ha tuttora posizione incognita. È però nota la direzione del vettore dC : spostandosi il centro di istantanea rotazione C lungo le due polari del moto, il vettore dC deve essere tangente a queste in C , cioè ortogonale al versore \mathbf{a} normale alle due polari in C (v. Figg. 2 e 3). Ne consegue che dC deve potersi esprimere come

$$dC = d\lambda \mathbf{k} \times \mathbf{a} \quad (10)$$

per un conveniente valore dello scalare infinitesimo $d\lambda$, dipendente da $d\vartheta$.

La sostituzione nell'Eq. (9) dell'espressione di dC fornita dall'Eq. (10) conduce a

$$\begin{cases} (r_0 - r_1) d\lambda = r_0 r_1 d\vartheta \\ (s_0 - s_1) (\mathbf{a} \cdot \mathbf{b}) d\lambda = s_0 s_1 d\vartheta \end{cases} \quad (11)$$

Il rapporto membro a membro delle due condizioni scalari (11) conduce infine la formula di Eulero-Savary

$$\frac{1}{r_0} - \frac{1}{r_1} = \left(\frac{1}{s_0} - \frac{1}{s_1} \right) \mathbf{a} \cdot \mathbf{b} \quad (12)$$

Per il caso particolare qui considerato come esempio, il centro di curvatura in M della curva involupata ε_0 (punto P_0 di Fig. 1) può essere determinato ricavando s_0 dall'Eq. (12) e sostituendo il valore trovato nella seconda delle Eq. (5), considerata per $j=0$ (P_0 coincide con P_0^* per $\vartheta = \bar{\vartheta}$).

Scelti i versi dei versori \mathbf{a} e \mathbf{b} lungo le rette passanti per O_0^* e O_1^* e, rispettivamente, P_0^* e P_1^* , il significato dei rimanenti simboli nell'Eq. (12) è definito dalle Eq. (6).

4. DISCUSSIONE

Si desidera ora giustificare brevemente l'ammissibilità dell'adozione del quadrilatero articolato ausiliario ai fini della dimostrazione del teorema di Eulero-Savary.

Come già affermato nella seconda sezione della presente memoria, i bilancieri $O_0^*O_1^*$ e $P_0^*P_1^*$ del quadrilatero articolato ausiliario vincolano i punti O_0^* , O_1^* e P_0^* , P_1^* a rimanere ad una distanza mutua rigorosamente costante durante il moto di π_1 rispetto a π_0 . Se però π_1 seguisse, rispetto a π_0 , la traiettoria di corpo rigido dettata dalle polari assegnate σ_0 e σ_1 , ciascuna di dette distanze sarebbe la somma di una quantità costante e di una quantità infinitesima del terzo ordine in $d\vartheta$.

Per consentire a π_1 di muoversi rispetto a π_0 esattamente come prescritto dal rotolamento delle due polari, si suppone ora che il quadrilatero articolato ausiliario abbia i bilancieri $O_0^*O_1^*$ e $P_0^*P_1^*$ di lunghezza variabile. I punti O_0^* e P_0^* sono immobili rispetto al piano fisso π_0 , mentre O_1^* e P_1^* sono immobili rispetto al piano mobile π_1 . Questo nuovo quadrilatero articolato è qui denominato quadrilatero articolato adattativo.

Il punto P_0^* , centro di curvatura in M della curva involupata ε_0 , è il centro di curvatura della traiettoria descritta da P_1^* rispetto al piano π_0 , nella posizione in cui viene a trovarsi P_1^* per $\vartheta = \bar{\vartheta}$. A sua volta P_1^* è un punto della biella del quadrilatero articolato adattativo.

Si consideri ora un generico punto H^* solidale con la biella del quadrilatero articolato adattativo. Il centro di curvatura della traiettoria descritta da H^* rispetto al piano fisso π_0 è determinabile conoscendo, oltre alla posizione di H^* , anche le derivate prima e seconda rispetto a ϑ del vettore posizione di H^* . La posizione di H^* rispetto a π_0 , funzione del parametro di configurazione ϑ , può essere espressa nel seguente modo

$$H^* = H^* [\vartheta, g(\vartheta)] \quad (13)$$

ove $g(\vartheta)$ esplicita la dipendenza di H^* dalle dimensioni cinematicamente rilevanti del quadrilatero articolato adattativo (i due bilancieri del quadrilatero articolato adattativo hanno lunghezza dipendente da ϑ ; nel quadrilatero articolato ausiliario, invece, tutti i lati hanno lunghezza costante e $g(\vartheta)$ sarebbe di fatto una quantità indipendente da ϑ).

Derivando l'Eq. (13) rispetto a ϑ si ottiene

$$\begin{aligned}\frac{dH^*}{d\vartheta} &= \frac{\partial H^*}{\partial \vartheta} + \frac{\partial H^*}{\partial g} \frac{dg}{d\vartheta} \\ \frac{d^2 H^*}{d\vartheta^2} &= \frac{\partial^2 H^*}{\partial \vartheta^2} + 2 \frac{\partial^2 H^*}{\partial \vartheta \partial g} \frac{dg}{d\vartheta} + \frac{\partial^2 H^*}{\partial g^2} \left(\frac{dg}{d\vartheta} \right)^2 + \frac{\partial H^*}{\partial g} \frac{d^2 g}{d\vartheta^2}\end{aligned}\tag{14}$$

Per $\vartheta = \bar{\vartheta}$ le derivate prima e seconda calcolate rispetto a ϑ delle lunghezze dei due bilancieri del quadrilatero articolato adattativo sono nulle (v. Eqq. (1) e (2)). Sono pertanto soddisfatte le seguenti due condizioni

$$\frac{dg}{d\vartheta} = \frac{d^2 g}{d\vartheta^2} = 0\tag{15}$$

Sostituendo le Eq. (15) nelle Eq. (14) si ottiene allora

$$\begin{aligned}\frac{dH^*}{d\vartheta} &= \frac{\partial H^*}{\partial \vartheta} \\ \frac{d^2 H^*}{d\vartheta^2} &= \frac{\partial^2 H^*}{\partial \vartheta^2}\end{aligned}\tag{16}$$

Queste condizioni mostrano che le derivate totali prima e seconda di H^* rispetto a ϑ sono uguali alle corrispondenti derivate parziali, come accadrebbe se le dimensioni cinematicamente rilevanti del quadrilatero articolato adattativo non dipendessero da ϑ o, in altri termini, come se il quadrilatero articolato adattativo fosse sostituito da quello ausiliario. Trova quindi giustificazione l'adozione del quadrilatero articolato ausiliario ai fini della determinazione del centro di curvatura di un generico punto di biella H^* e, in particolare, del punto P_1^* .

5. CONCLUSIONI

Grazie all'espedito del quadrilatero articolato ausiliario, la dimostrazione qui proposta del teorema di Eulero-Savary nel caso piano potrebbe risultare più comprensibile di altre a chi si cimenta per la prima volta nello studio di questo argomento di cinematica.

La strategia di dimostrazione adottata nel caso piano è facilmente estensibile al caso sferico mediante il ricorso ad un quadrilatero articolato ausiliario sferico.

BIBLIOGRAFIA

- [1] Levi-Civita, T., Amaldi, U., 1974, *Lezioni di Meccanica razionale*, Vol. 1, Zanichelli, Bologna, pp. 251-256.
- [2] Di Benedetto, A., Pennestrì, E., 1993, *Introduzione alla cinematica dei meccanismi*, Vol. 2, C.E.A., Milano, pp. 117-127.
- [3] Hunt, K.H., 1978, *Kinematic Geometry of Mechanisms*, Clarendon Press, Oxford, pp. 123-126.
- [4] Bottema, O., Roth, B., 1979, *Theoretical Kinematics*, North-Holland, Amsterdam, pp. 276-278.
- [5] Prosciutto, A., 1956, *Elementi di Meccanica applicata alle macchine*, 5^a edizione, Pàtron, Bologna, pp. 37-39.
- [6] Hartenberg, R.S., Denavit, J., 1964, *Kinematic Synthesis of Linkages*, McGraw-Hill, New York, pp. 194-201.
- [7] Paul, B., 1979, *Kinematics and Dynamics of Planar Machinery*, Prentice-Hall, New Jersey, pp. 199-201.
- [8] Funaioli, E., Maggiore, A., Meneghetti, U., 1987, *Lezioni di Meccanica applicata alle macchine*, Vol. 1, Pàtron, Bologna, pp. 130-131.

A NEW PROCEDURE FOR THE OPTIMIZATION OF A DIELECTRIC ELASTOMER ACTUATOR

Rocco Vertechy
PERCRO Laboratory,
Scuola Superiore Sant'Anna, Pisa, Italy
E-mail: r.vertechy@sssup.it

Giovanni Berselli
Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: giovanni.berselli@unibo.it

Vincenzo Parenti Castelli
Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: vincenzo.parenti@unibo.it

Gabriele Vassura
Department of Mechanical and Aeronautical Engineering,
University of Bologna, Italy
E-mail: gabriele.vassura@unibo.it

Abstract. *A novel mathematical procedure is proposed, which makes it possible to optimize lozenge-shaped dielectric-elastomer-based linear actuators for known materials and desired force/stroke requirements. Simulation results are provided which both demonstrate the efficacy of the novel procedure with respect to traditional design approaches and show that simpler, cheaper, lighter and better-behaved lozenge-shaped actuators can be conceived which do not require any integration of compliant frame elements.*

Keywords: *dielectric elastomers, smart actuators, optimal design*

1. INTRODUCTION

Dielectric Elastomers (DE) are deformable dielectrics which can experience deviatoric (isochoric) finite deformations in response to applied large electric fields while, at the same time, alter the applied electric fields in response to the deformations undergone [1-4]. Thanks to the strong electro-mechanical coupling, DE intrinsically offer great potentialities for conceiving novel solid-state mechatronic devices, such as sensors and actuators, which are more integrated, lightweight, economic, silent, resilient and disposable than equivalent systems based on traditional technologies [5, 6].

For actuation usage, DE are usually shaped in thin films coated with compliant electrodes on both sides and piled one on the other to form an Electro-Active Film (EAF) [2-4]. Activation of the EAF via the placement of differential electric potentials (hereafter

also called voltages) between the electrodes can induce film area expansions and, thus, point's displacements which can be used to produce useful mechanical work (whenever forces are applied to such points).

Usually, DE-based actuators are obtained by first uniformly pre-stretching the EAF (which is necessary since the film has negligible flexural rigidity) and then by coupling some segment of its boundary to some portion of a flexible supporting frame, either an elastic structural element [7, 8] (for instance a helical spring) or a compliant mechanism [9-13] (for instance a four-bar mechanism with elastic revolute joints). Besides preventing the development of current arcs around the EAF border, protecting the EAF edges and providing a firm support for the application of external forces, the major roles of the flexible polymeric frame are: 1) to coerce the EAF expansion in preferred directions; 2) to maintain the EAF in a tensioned state so as to prevent buckling effects; and 3) to provide a desired actuator stiffness characteristic (usually a nearly naught stiffness through the whole actuator stroke).

Practical examples of DE-based linear actuators are depicted in Figures 1.a and 1.b. The actuators are based on a lozenge-shaped pre-stretched EAF coupled to elastic frames made by a four-bar mechanism with equal links and either two identical non-linear tension springs (Fig. 1.a, [9]) or a compliant symmetric double slider-crank mechanism with elastic joints on the slider pivots (Fig. 1.b, [12]). As depicted in Fig. 2, in such actuators, activation of the EAF makes it possible to control the relative distance x (hereafter also called "EAF length" or "actuator length") of the centers O and P of two opposing revolute pairs of the four-bar mechanism, which are the points of application of the external forces F .

The actual procedure for designing DE-based actuators consists in first selecting the appropriate EAF topology; second, empirical determining EAF pre-stretches and dimensions; and, finally, mathematical or experimental designing a flexible polymeric frame which complies with points 1-3 described above. Although valid, this procedure is not rigorous, neither provides optimal actuators.

In this paper a novel procedure is proposed for the optimal design of lozenge-shaped DE-based linear actuators like the ones depicted in Fig. 2. The procedure is efficient, based on a mathematical model and makes it possible to improve the actuator force/stroke performance while, at the same time, eliminating the need of supporting frame compliant elements (which simplify the DE-based actuator drastically). In the following, first the general optimization method is presented, which is valid for wide-ranging DE materials

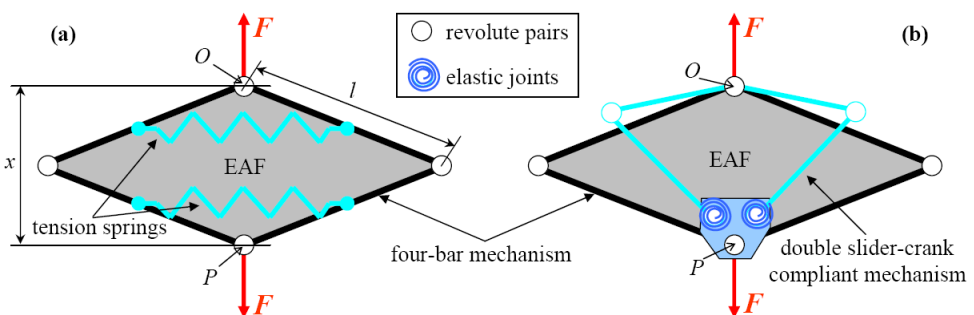


Figure 1. Lozenge-shaped DE-based linear actuators

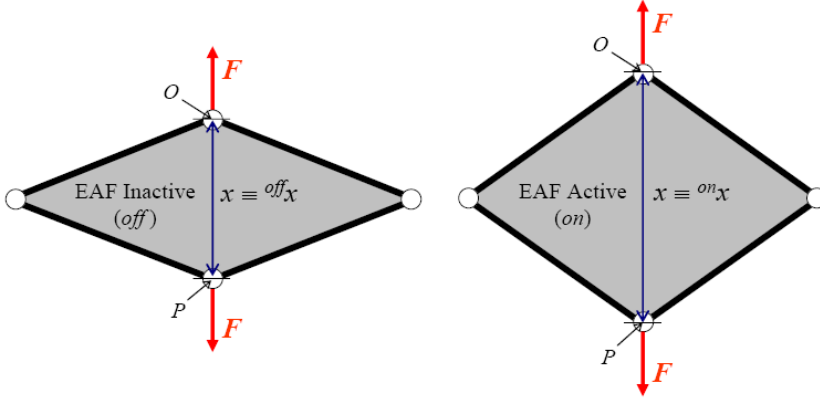


Figure 2. Inactive (*off*) and active (*on*) lozenge-shaped DE-based linear actuator

and actuator force/stroke requirements, then a case study is provided that highlights both the efficiency of the procedure and the improved performances of the resulting actuator.

2. GENERAL ASPECTS OF DE-BASED ACTUATOR DESIGN

By definition [2-4], DE are incompressible, hyper-elastic linear dielectrics whose electrical polarization is fairly independent of material deformation [4]. For such materials, it is well known [2] that EAF activation generates an electric field, $E = V/t$ (V being the activation voltage applied between the EAF electrodes and t is the actual thickness of the DE film amid the EAF electrodes), and an electrically-induced Cauchy stress, $\sigma_{em} = \varepsilon E^2$ (ε being the DE electric permittivity), both acting in the DE film thickness direction. Consequently, the mechanical stress field in a stretched and activated single-layered EAF, which is free to deform in its thickness direction, is given by the following relationships [14]:

$$\sigma_1 = \lambda_1 \frac{\partial \Psi}{\partial \lambda_1} - \sigma_{em}, \quad (1.1)$$

$$\sigma_2 = \lambda_2 \frac{\partial \Psi}{\partial \lambda_2} - \sigma_{em}, \quad (1.2)$$

$$\sigma_3 = -\sigma_{em} = -\varepsilon \frac{V^2}{\lambda_3^2 t_0^2}, \quad (1.3)$$

$$\lambda_3 = 1/\lambda_1 \lambda_2, \quad (1.4)$$

where λ_i and σ_i ($i = 1, 2, 3$) are, respectively, the principal stretches and Cauchy stresses (the third principal direction coinciding with the film thickness direction), $\Psi = \Psi(\lambda_1, \lambda_2)$ is

the incompressible DE strain-energy function [14] and t_0 is the unstretched DE film thickness.

By construction, when coupled with a four-bar mechanism having links of equal length, lozenge-shaped EAFs expand uniformly without changing their edge length l and principal stretch/stress directions [10]. Thus, their deformation state is characterized by the following principal stretches

$$\lambda_1 = \lambda_{1,p} x/x_p, \quad (2.1)$$

$$\lambda_2 = \lambda_{2,p} \sqrt{(4l^2 - x^2)/(4l^2 - x_p^2)}, \quad (2.2)$$

where, referring to Figures 1 and 2, $\lambda_{1,p}$ and $\lambda_{2,p}$ are the EAF principal pre-stretches in the configuration where $x \equiv x_p$ (x_p can be chosen as desired). Of course, the first and second principal stretch/stress directions are respectively aligned and orthogonal to the line joining the points O and P .

Considering Fig. 3, the external forces (either force off, ^{off}F , or force on, ^{on}F , depending on the EAF activation state) that must be supplied at O and P , and directed along the line joining these points, to balance the EAF stress field are

$$^{off}F(x) = \frac{t_0}{2} \left(\frac{\sqrt{4l^2 - x_p^2}}{\lambda_{2,p}} \frac{\partial \Psi}{\partial \lambda_1} - \frac{xx_p}{\lambda_{1,p} \sqrt{(4l^2 - x^2)}} \frac{\partial \Psi}{\partial \lambda_2} \right), \quad (3.1)$$

$$^{on}F(x, V) = ^{off}F(x) + ^{em}F(x, V), \quad (3.2)$$

$$^{em}F(x, V) = - \frac{\varepsilon V^2 \lambda_{1,p} \lambda_{2,p}}{t_0} \frac{x(2l^2 - x^2)}{x_p \sqrt{4l^2 - x_p^2}}, \quad (3.3)$$

which, by convention, are positive if directed according to the arrows depicted in Figures 1, 2 and 3.

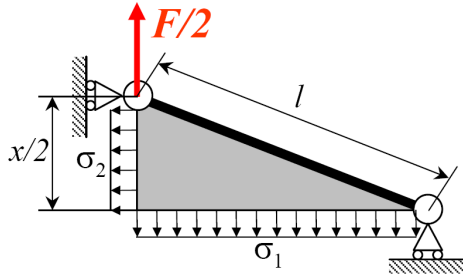


Figure 3. Force equilibrium of lozenge-shaped EAF

Figure 3 represents the EAF loading scheme reduced to a quarter of the lozenge-shaped EAF because of symmetry. Note that ${}^{em}F(x, V)$, given by Eq. (3.3), is an “electric” force due to EAF activation.

The mathematical Force-Length (FL) curves, ${}^{off}F(x)$ and ${}^{on}F(x, V)$, obtained by Eqs. (1-3) for different values of activation voltages (respectively $V = 0\text{V}$ and $V = 2.2\text{kV}$, $V = 6.5\text{kV}$ and $V = 8.7\text{kV}$) are plotted in Fig. 4 for a typical lozenge shaped EAF, composed of a single acrylic DE (3M-VHB 4905) and two silver conductive grease electrodes (CircuitWorks-CW7100). The film pre-stretches are $\lambda_{1,p} = 2.2$ and $\lambda_{2,p} = 5$ (as recommended in [9, 10]); the film dimensions are $t_0 = 1.5\text{ mm}$, $l = 70\text{mm}$ and $x_p = 15\text{mm}$. The mathematical curves are obtained by using a first-order Ogden type [15] strain-energy function

$$\Psi = \frac{\mu}{\alpha} \left(\lambda_1^\alpha + \lambda_2^\alpha + \lambda_1^{-\alpha} \lambda_2^{-\alpha} - 3 \right), \quad (4)$$

with parameters $\mu = 60\text{kPa}$ and $\alpha = 1.8$, and a DE electric permittivity $\varepsilon = 4.5 \cdot 8.85 \cdot 10^{-12} \text{F/m}$ [10]. The figure shows that the EAF curves ${}^{off}F(x)$ and ${}^{on}F(x, V)$ are usually nonparallel and nonlinear. This complicates the DE-based actuator design and control. Indeed, a well-behaved actuator should feature linear (typically with zero slope) FL curves and available trust $T(x)$, $T(x) = [{}^{off}F(x) - {}^{on}F(x, V_{max})]$ (V_{max} being the maximum operating voltage), independent of actuator length, x .

To face this issue, traditional design procedures propose to couple EAF (of experimentally predefined pre-stretch and geometry) with suitably sized elastic frames made of polymeric materials [9-13]. That is, the actual design procedure for lozenge-shaped DE actuators consists in the following sequential steps: i) empirical determination of film pre-stretches which guarantee the maximal area expansion of the activated EAF [9-11]; ii) experimental determination of EAF geometrical dimensions which satisfy given actuator force/stroke requirements [9-11]; iii) mathematical or experimental design of the elastic frame [9-12]. Although valid, this design approach is not rigorous, neither provides optimal actuators. Indeed, elastic frames can only linearize either the ${}^{off}F(x)$ or ${}^{on}F(x, V)$ EAF curves. Moreover, by introducing more components, the resulting actuators become bulkier, heavier, more expensive and complex, and are likely to show degraded efficiency, reliability and repeatability.

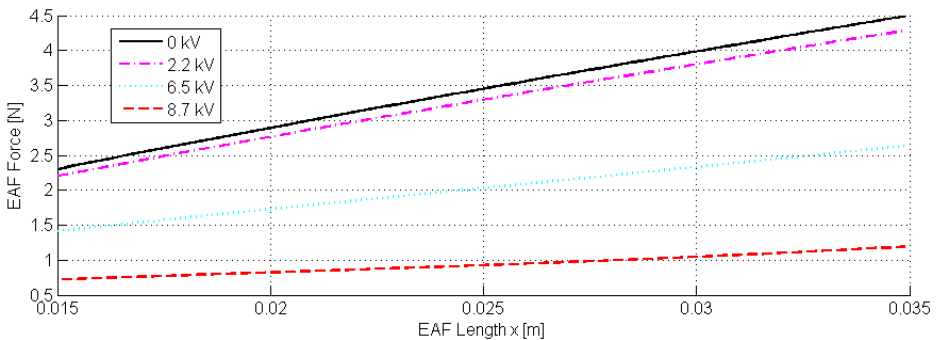


Figure 4. Force-Length (FL) curves of a typical lozenge-shaped EAF

3. THE NOVEL DESIGN APPROACH

In this section, a mathematical procedure is described that makes it possible to design well-behaved lozenge-shaped DE-based actuators, which do not need to be coupled with properly dimensioned compliant frames.

Consider the adimensional parameter $\chi = x/l$ which uniquely identifies the lozenge configuration. Differently from the actual design procedure, the proposed approach has the following sequential steps: i) given the desired initial and final actuator lengths, x_i and x_f (and thus the ratio $\delta = x_i/x_f$), identify the range of lozenge configurations, $[\delta\chi_f, \chi_f]$ ($\chi_f = x_f/l$ being the final actuator configuration), that minimize the maximum actuator trust relative error \underline{e}_T , $\underline{e}_T = [\max(T(x))/\min(T(x)) - 1]$ within $x_i \leq x \leq x_f$; ii) find the lozenge geometry which satisfy the desired actuator lengths, x_f and x_i , and the desired force trust \underline{T} , $\underline{T} = \min[T(x)]$ within $x_i \leq x \leq x_f$, for given maximum bearable (limiting) electric field E_i ; iii) find the EAF pre-stretches which, besides preventing film failures (i.e. buckling or rupture), guarantee the desired linear FL curve, ^{off}F , of the actuator in its inactive state.

Mathematical Insights On The EAF Force

Posing $E_l = E_f = V/t_f$ (t_f being the DE film thickness in the final actuator configuration) and, momentarily, $\chi_p = \chi_f$, equations (3.1) and (3.3) can be rewritten as

$$^{off}F = t_f \chi_f \sqrt{1 - \chi_f^2/4} \ ^{off}f(\chi, \lambda_{1,p}, \lambda_{2,p}), \quad (5.1)$$

$$^{em}F = \frac{t_f \varepsilon E_l^2}{\chi_f \sqrt{4 - \chi_f^2}} \ ^{em}f(\chi), \quad (5.2)$$

where

$$^{off}f = \frac{\lambda_1}{\chi} \frac{\partial \Psi}{\partial \lambda_1} - \frac{\lambda_2 \chi}{4 - \chi^2} \frac{\partial \Psi}{\partial \lambda_2}, \quad (5.3)$$

$$^{em}f = \chi (\chi^2 - 2), \quad (5.4)$$

which show that the trust relative error, $(T(x)/\underline{T} - 1)$, depends on χ (and thus on $\delta = x_i/x_f$) only, while the nonlinearity of $^{off}F(x)$, and thus of $^{on}F(x, V)$, depends on χ , $\lambda_{1,p}$, $\lambda_{2,p}$ and on the DE mechanical properties hidden in the strain-energy function Ψ [14].

A plot of the adimensional “electric” force, ^{em}f , is given in Fig. 5 which highlights that the EAF can sustain significant compressive external forces only if $\chi_f \leq 2^{0.5}$ and provide a quasi constant trust $T(x)$ if $\delta\chi_f$ and χ_f are close to 0.8. Therefore, in this paper $\chi_f \leq 2^{0.5}$ is considered only. Note that if $\chi_f \leq 2^{0.5}$, the maximum electric field E , which acts on the DE for a given electric potential V , is at χ_f (i.e. the actuator configuration for which EAF area and thickness are maximum and minimum respectively), for this reason $E_l = E_f$ has been posed in Eq. (5).

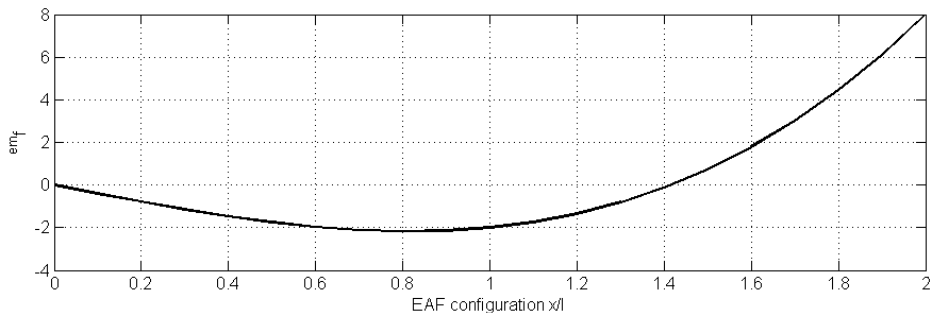


Figure 5. Adimensional electric force ^{em}f vs EAF configuration x/l

The Optimization Procedure

Following the considerations highlighted in the previous sub-section, actuator optimization can be conducted according to the following sequential steps:

1) Given the desired initial and final actuator lengths, x_i and x_f (and thus $\delta = x_i/x_f$), then find χ_f that minimizes \underline{e}_r , i.e.

$$\chi_f = \sqrt{\frac{2}{\delta^2 + \delta + 1}}, \quad (6.1)$$

the resulting error being

$$\underline{e}_r = \frac{\sqrt{8/3} - (2/3)^{3/2}}{2\chi_f - \chi_f^3} - 1; \quad (6.2)$$

2) Given χ_f and the desired final actuator length x_f solve for the lozenge length $l = x_f/\chi_f$;

3) Given χ_f , l , the DE electrical properties, ε and E_i , and the desired trust \underline{T} find t_f as

$$t_f = \frac{\chi_f \sqrt{4 - \chi_f^2 \underline{T}}}{^{em}f(\chi_f) l \varepsilon E_i^2}, \quad (7)$$

4) Given χ_f , l , t_f , the DE mechanical properties and a desired force profile $^{off}\underline{F}(x)$, find $\underline{\lambda}_{1,p}$ and $\underline{\lambda}_{2,p}$ that minimize

$$\mathfrak{Z}(\lambda_{1,p}, \lambda_{2,p}) = \int_{x_i}^{x_f} \left[^{off}F(x, \lambda_{1,p}, \lambda_{2,p}) - ^{off}\underline{F}(x) \right]^2 dx, \quad (8.1)$$

subjected to the following anti-failure criteria

$$\sigma_1(x, \underline{\lambda}_{1,p}, \underline{\lambda}_{2,p}) > 0, \forall x \in [x_i, x_f], \quad (8.2)$$

$$\sigma_2(x, \underline{\lambda}_{1,p}, \underline{\lambda}_{2,p}) > 0, \forall x \in [x_i, x_f], \quad (8.3)$$

$$\Phi(\lambda_1(x), \lambda_2(x)) < 0, \forall x \in [x_i, x_f], \quad (8.4)$$

where the relations (8.2) and (8.3) are derived from equations (1.1) and (1.2) and guarantee that the EAF will not buckle in the desired actuator range of motion, and Φ is a suitable fracture criterion of the DE [16];

5) Given t_f , $\underline{\lambda}_{1,p}$ and $\underline{\lambda}_{2,p}$ find the unstretched film thickness

$$t_0 = t_f \underline{\lambda}_{1,p} \underline{\lambda}_{2,p}. \quad (9)$$

4. CASE STUDY

The procedure presented in Section 3 has been used for designing a linear actuator. The actuator position requirements are $x_i = 15\text{mm}$ and $x_f = 35\text{mm}$ (thus $\delta = 0.43$); the actuator force requirements are a tensile force $^{off}\underline{F}(x) = +0.5\text{N}$, a compressive force $^{on}\underline{F}(x, V_{max}) = -1\text{N}$, and thus an available trust $\underline{T} = 1.5\text{N}$. The DE and electrode materials are the same as those used for the example examined in Fig. 4. That is, an acrylic DE (3M-VHB 4905) and silver conductive grease electrodes (CircuitWorks-CW7100) such that $\varepsilon = 4.5 \cdot 8.85e^{-12}\text{F/m}$, $\mu = 60\text{kPa}$ and $\alpha = 1.8$. The Kawabata's criterion [16]

$$\Phi(\lambda_1(x), \lambda_2(x)) = \max_{x \in [x_i, x_f]} [\lambda_1(x), \lambda_2(x)] < 9, \quad (10)$$

is used to check for DE mechanical rupture. To avoid DE electric breakdown, the limiting electric field is set to $E_f = 150\text{MV/m}$.

The resulting actuator geometrical dimensions are $\chi_f = 1.113$, $l = 31\text{mm}$, $\underline{\lambda}_{1,p} = 2.12$ and $\underline{\lambda}_{2,p} = 8.18$ (referred to $x_p = 15\text{mm}$), and $t_0 = 4\text{mm}$. For safety reasons, limiting the required electric potential V below 10kV , the resulting EAF is subdivided in two layers each having $t_0 = 2\text{mm}$. Note that, for the given DE material, these design parameters are obtained analytically. The mathematical Force-Length (FL) curves, $^{off}\underline{F}(x)$ and $^{on}\underline{F}(x, V)$, of the optimized actuator are plotted in Fig. 6 for different values of activation voltages (respectively $V = 0\text{V}$ and $V = 2.2\text{kV}$, $V = 6.5\text{kV}$ and $V = 8.7\text{kV}$). As shown, the EAF fully satisfy the desired actuator force/stroke requirements. The FL curves are nearly constant independently of the activation voltages. The maximum (at $V_{max} = 8.7\text{kV}$) trust relative error is $\underline{e}_T = 28\%$ (of $\underline{T} = 1.5\text{N}$), which is slightly less than that calculated via Eq. (6.2).

Comparison between Figures 4 and 6 highlights that the optimized EAF performs better than the typical EAF designed according to the conventional design procedure. Indeed, for the latter, the FL curves are non-parallel, with slope depending on the activation voltage, and the maximum (at $V_{max} = 8.7\text{kV}$) trust relative error is $\underline{e}_T = 108\%$ (of $\underline{T} = 1.58\text{N}$). Moreover, while the optimized EAF can simply be coupled to a four-bar mechanism, in order to satisfy the desired $^{off}\underline{F}(x) = 0.5\text{N}$ the typical EAF requires a compliant frame which, besides bringing design, manufacturing, reliability and efficiency

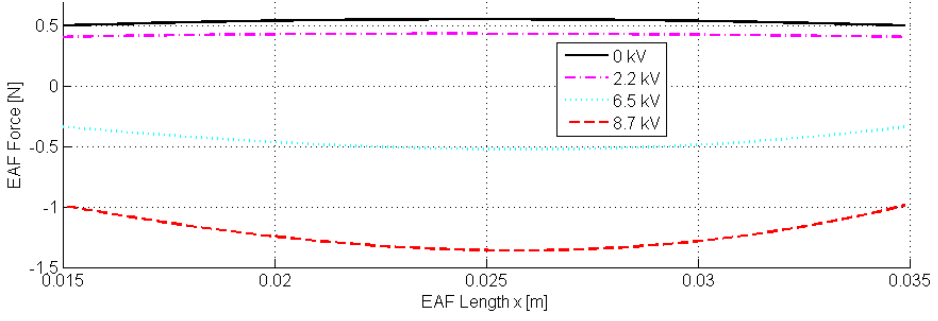


Figure 6. Force-Length (FL) curves of the optimized lozenge-shaped EAF

issues, is likely to worsen \underline{e}_T . The optimized EAF is also characterized by larger ratios between required trust, \underline{T} , and necessary DE volume, ν , and smaller ratios between maximum actuator capacitance, C_f , and desired actuator trust, \underline{T} , and between actuator capacitance variation, $\Delta C = C_f - C_i$, and desired actuator trust, \underline{T} . Indeed, for equal V , E_l , ε , x_f , δ , the abovementioned ratios can be written as

$$\frac{\underline{T}}{\nu} = \frac{\varepsilon E_l^2}{x_f} \frac{2\delta(2 - \delta^2 \chi_f^2)}{4 - \chi_f^2}, \quad (11.1)$$

$$\frac{C_f}{\underline{T}} = \frac{x_f}{V^2} \frac{4 - \chi_f^2}{2\delta(2 - \delta^2 \chi_f^2)}, \quad (11.2)$$

$$\frac{\Delta C}{\underline{T}} = \frac{x_f}{V^2} \frac{4 - \chi_f^2 - \delta^2(4 - \delta^2 \chi_f^2)}{2\delta(2 - \delta^2 \chi_f^2)}, \quad (11.3)$$

which, by plotting in Figures 7, 8 and 9 the adimensional quantities $\underline{T}^* = 2\delta(2 - \delta^2 \chi_f^2)/(4 - \delta^2 \chi_f^2)$, $C_f^* = (4 - \chi_f^2)/[2\delta(2 - \delta^2 \chi_f^2)]$ and $\Delta C^* = [4 - \chi_f^2 - \delta^2(4 - \delta^2 \chi_f^2)]/[2\delta(2 - \delta^2 \chi_f^2)]$ respectively versus EAF final configuration, χ_f , show that, for different δ , the ratio \underline{T}/ν increases as χ_f increases and that both the ratios C_f/\underline{T} and $\Delta C/\underline{T}$ decrease as χ_f increases (the optimized and the typical actuators are characterized by $\chi_f = 1.113$ and $\chi_f = 0.5$ respectively). This means that the optimized EAF requires less DE material for producing the same actuator trust of the typical EAF and that the electric time constant of the resulting optimized actuator will be smaller and less variable than that of the typical actuator (i.e. the optimized actuator will have better dynamic response, and will be easier to drive and control than the actuator designed according to the traditional procedure). Moreover, the optimized EAF has lower lateral encumbrance than the typical EAF since the lengths of the lozenge edges are 31mm and 70mm respectively.

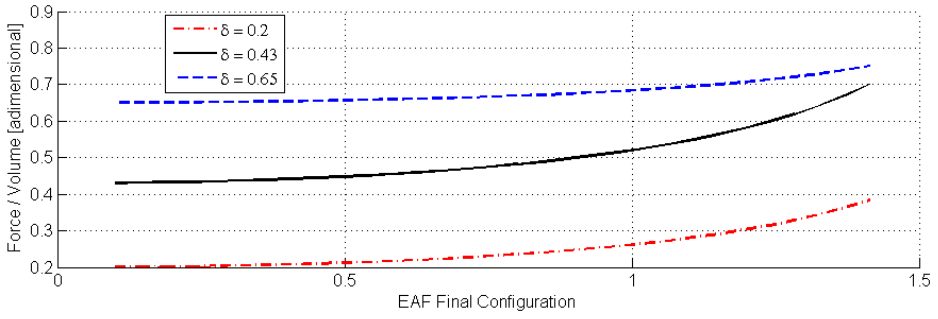


Figure 7. Adimensional ratio \underline{T}^*

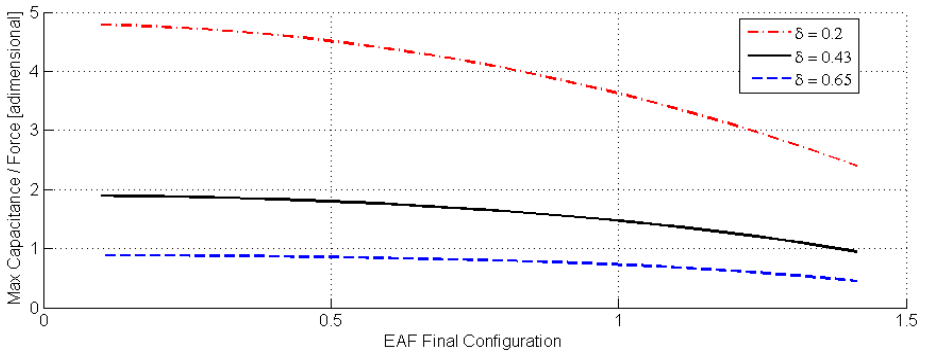


Figure 8. Adimensional ratio C_f^*

Finally, the abovementioned results show that the empirical condition, $7 \leq \underline{\lambda}_{1,p} \cdot \underline{\lambda}_{2,p} \leq 11$ that maximizes the free area expansion of EAF, which was suggested in [9, 10] for choosing the film pre-stretches at x_i is not a necessary one for lozenge-shaped DE-based linear actuators.

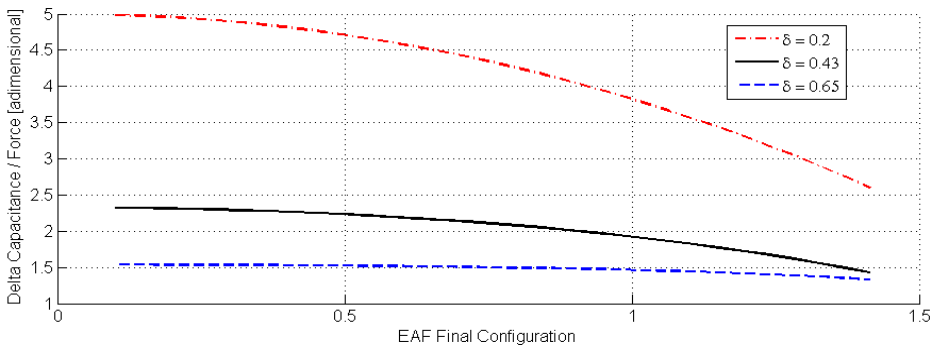


Figure 9. Adimensional ratio ΔC^*

5. CONCLUSIONS

A novel mathematically based procedure for the optimization of lozenge-shaped dielectric-elastomer-based linear actuators has been presented. The procedure has been applied to a case study. Simulation results showed that traditional design approaches are not rigorous and neither provides actuators with optimal performances, while the novel procedure makes it possible to develop improved actuators having elastic-less frames, maximum force density and reduced electrical activation requirements.

REFERENCES

- [1] Toupin, R.A., "The Elastic Dielectrics", *J. Rational Mech. Anal.*, **5**, pp. 849-915, 1956
- [2] Pelrine, R.E., Kornbluh, R.D., Joseph, J., 1998. "Electrostriction of polymer dielectrics with compliant electrodes as a means of actuation", *Sensors Actuators A*, **64**, pp. 77-85.
- [3] Pelrine, R., Kornbluh, R., Kofod, G., 2000. "High-Strain Actuator Materials Based on Dielectric Elastomers", *Advanced Materials*, **12**(16), pp. 1223-1225.
- [4] Kofod, G., Sommer-Larsen, P., Kornbluh, R., Pelrine, R., 2003. "Actuation Response of Polyacrylate Dielectric Elastomers". *Journal of Intelligent Material Systems and Structures*, **14**, pp. 787-793.
- [5] Bar-Cohen, J., 2001. *Electroactive Polymer (EAP) Actuators as Artificial Muscles: Reality, Potential and Challenges*. Bellingham, WA, SPIE Press.
- [6] Kim, L.K., Tadokoro, S., 2007. *Electroactive Polymers for robotic applications*, Springer.
- [7] Kofod, G., Paajanen, M., Bauer, S., 2006. "Self-organized minimum-energy structures for dielectric elastomer actuators". *Applied Physics A: Materials Science & Processing*, **85**(2), pp. 141-143.
- [8] Pei, Q., Pelrine, R., Stanford, S., Kornbluh, R., Rosenthal, M., 2003. "Electroelastomer rolls and their application for biomimetic walking robots". *Synth. Met.* **135/136**, pp. 129-31.
- [9] Vogan, J., 2004. "Development of Dielectric Elastomer Actuators for MRI Devices". *M.S. Thesis*, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- [10] Plante, J.S., 2006. "Dielectric elastomer actuators for binary robotics and mechatronics". *Ph.D. Thesis*, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA.
- [11] Wingert, A., Lichter, M.D., Dubowsky, S., 2006. "On the design of large degree-of-freedom digital mechatronic devices based on bistable dielectric elastomer actuators". *IEEE/ASME Transactions on Mechatronics*, **11**(4), pp. 448-456.
- [12] Berselli, G., Vertechy, R., Vassura, G., Parenti Castelli, V., 2008. "A compound-structure frame for improving the performance of a Dielectric Elastomer Actuator". *Advances in Robot Kinematics: Analysis and Design*, Springer Netherlands, ISBN: 978-1-4020-8599-4, pp. 291-299.
- [13] Berselli, G., Vertechy, R., Vassura, G., Parenti Castelli, V., 2008. "Design of a single-acting constant-force actuator based on dielectric elastomers". *ASME IDETC/CIE 2008 International Design Engineering Technical Conferences*, New York City, USA.
- [14] Holzapfel, G.A., 2001. *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*, John Wiley & Sons.

- [15] Ogden, R.W., 1972. "Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids". *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, **326**(1567), pp. 565-584.
- [16] Hamdi, A., Nait Abdelaziz, M., Ait Hocine, N., Heuillet, P., Benseddiq, N., 2006. "A fracture criterion of rubber-like materials under plane stress conditions". *Polymer Testing*, **25**(8), pp. 994-1005

SULL'INNESCO DELLA FRATTURA IN CONTATTI NON HERTZIANI

Angela Benedetti

E-mail: ang.benedetti@libero.it

Pier Gabriele Molari

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Aeronautiche e di Metallurgia.

Università di Bologna

E-mail: piergabriele.molari@unibo.it

Piero Morelli

Dipartimento di Ingegneria delle Costruzioni Meccaniche, Aeronautiche e di Metallurgia.

Università di Bologna

E-mail: piero.morelli@unibo.it

Sommario. *In questo lavoro si rivisita e si discute uno studio numerico-sperimentale sulla fatica di contatto, condotto in condizioni di carico che comportino deformazioni plastiche del materiale. Tale studio è stato presentato in parte al VI congresso “Fracture and Damage Mechanics” ad Harbin (Cina) nel 2006 [1]. Si definiscono le geometrie e le proprietà meccaniche dei corpi, e la tipologia del carico affaticante di contatto. Si mostrano i risultati di uno studio agli elementi finiti dello stato tensionale, e si discutono le differenze che emergono simulando i legami costitutivi dei materiali con andamenti lineari piuttosto che multi-lineari. Si mettono in relazione le zone delle mappe tensionali caratterizzate dagli stati di sollecitazione più gravosi, con i punti di innesco delle cricche e con le direzioni di propagazione delle fratture osservati sperimentalmente.*

Parole chiave: *meccanica del contatto, fatica dei materiali metallici, innesco della frattura, propagazione della frattura, leghe di alluminio.*

1. INTRODUZIONE

Lo studio del danneggiamento per fatica, che si verifica nelle zone di contatto fra i componenti di un sistema meccanico, è determinante sia per la valutazione della capacità delle superfici di continuare a trasmettere i carichi di esercizio in modo funzionalmente corretto, sia per identificare gli effetti che le fratture potrebbero produrre nei confronti dell'affidabilità e della sicurezza del sistema nel suo complesso.

Sono note, in letteratura, le distribuzioni dei campi di deformazione e di tensione che si producono localmente ai contatti, per diverse configurazioni geometriche, nell'ipotesi che il legame tra la tensione e la deformazione del materiale sia caratterizzato dalla relazione di proporzionalità lineare descritta dalla legge di Hooke [2,3]. In queste condizioni, e nei limiti nei quali le dimensioni dell'area dell'impronta possano essere considerate trascurabili

rispetto alle dimensioni dei corpi, i contatti soddisfano le ipotesi formulate da Hertz nel noto articolo del 1882, e possono dare luogo sia a fratture superficiali, che a fratture sub-superficiali.

Le prime si innescano in corrispondenza del bordo dell'impronta, ove agiscono tensioni radiali di trazione [4,5]. I punti di innesco possono essere molteplici. A seguito dell'accrescimento, che avviene in direzione perpendicolare alle tensioni di innesco, le fratture di superficie si uniscono distribuendosi lungo una o più ellissi, omotetiche alla forma dell'impronta e centrate sull'asse del contatto. In un secondo tempo, le cricche diffondono dalla superficie verso l'interno del materiale, incurvandosi nella direzione del centro del contatto.

La seconda tipologia di fratture, invece, si innesca in corrispondenza dell'asse dell'impronta, ad una certa profondità dalla superficie, ove la tensione di taglio è massima [4,5]. In ragione del modo con il quale viene applicato il carico affaticante, il fronte di tali fratture può propagarsi sotto pelle, oppure dirigersi verso la superficie del corpo [6], dando così luogo all'asportazione (pitting) di materiale. Nel secondo caso, i detriti che si formano possono concorrere ad incrementare la velocità di danneggiamento delle superfici in contatto, innescando fenomeni di usura di tipo abrasivo.

Quando l'intensità dei carichi di contatto è tale da comportare deformazioni plastiche nel materiale di uno o di entrambi i corpi nell'intorno della zona di contatto, le ipotesi di Hertz non sono più soddisfatte, e non è quindi possibile sfruttare i risultati noti per studiare i campi deformativo e tensionale locali al contatto. In queste condizioni sono state effettuate alcune prove sperimentali di fatica assiale di tipo pulsante [1], azionando un indentatore a testa sferica, in acciaio temprato (si veda il pezzo nr. 5 dell'assieme di Fig. 1), contro la faccia piana di provette cilindriche in lega di alluminio AA6082 (pezzo nr. 2 di Fig. 1). In queste ultime è stato osservato un terzo modo di innesco e di propagazione della frattura, che viene messo a confronto con i risultati di uno studio numerico dello stato tensionale nel contatto, realizzato tenendo conto delle caratteristiche elasto-plastiche dei materiali.

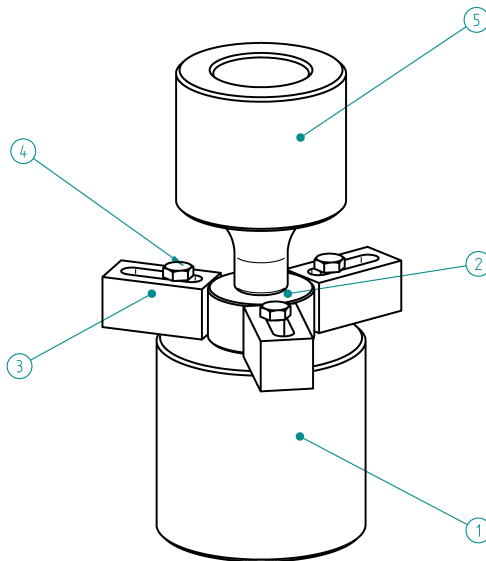


Figura 1. Assieme dell'attrezzatura.

2. ATTREZZATURA SPERIMENTALE

Il contatto oggetto dello studio è stato realizzato tra superfici a curvatura non conforme, del tipo sfera su piano. Essendo la geometria piana di più semplice esecuzione rispetto a quella sferica, si è deciso di indurre la frattura sul corpo piano, per il quale è stato quindi scelto un materiale più duttile e meno resistente rispetto a quello impiegato per la sfera. Sono state utilizzate, quali piani di contatto, le superfici lucidate di testa di provette cilindriche in AA6082, quotate in Fig. 2.

Il corpo sferico, in acciaio 38NiCrMo4, è stato lavorato come in Fig. 3, in modo da essere collegato alla traversa strumentata di una macchina di prova servoidraulica Instron, con 100 kN di capacità massima di carico. La testa a curvatura sferica è stata temprata e lucidata.

Le prove sono state condotte in compressione, con rapporto di carico $R=|\sigma_{min}|/|\sigma_{max}|=0.1$, a frequenza di 60 Hz. La temperatura di prova nell'ambiente è stata di 20 °C, e tra le superfici a contatto non è stato interposto alcun tipo di lubrificante.

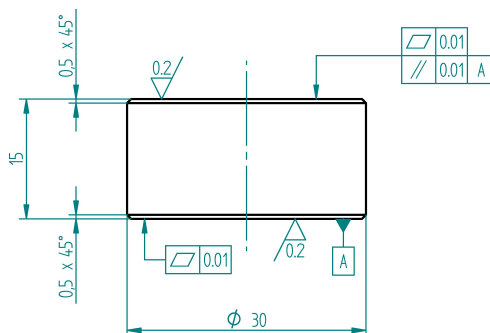


Figura 2. Provetta in AA6082.

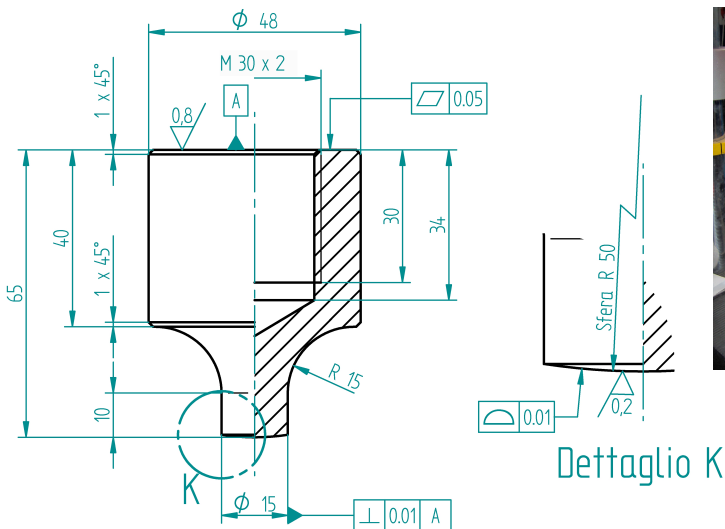


Figura 3. Indentatore a testa sferica in acciaio 38NiCrMo4. Attrezzatura montata.

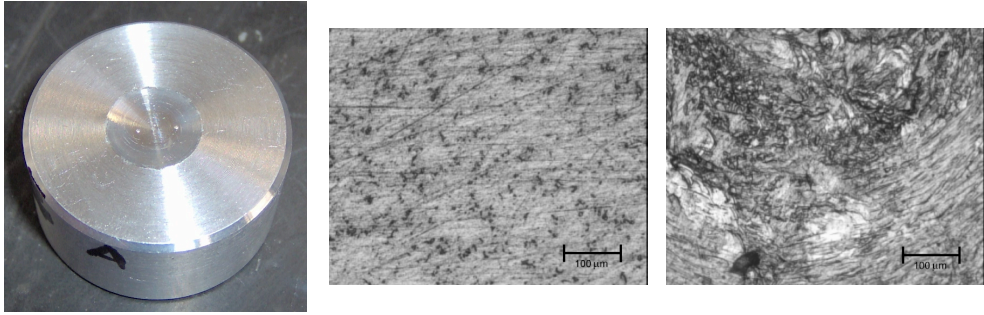


Figura 4. Impronta sferica, stato superficiale prima e dopo la prova.

3. MICROGRAFIE DELLE FRATTURE

Deformazione plastica superficiale

Si mostra, nella prima fotografia di Fig. 4, l'estensione della deformazione plastica, in forma di concavità sferica di diametro pari a 13 mm e profondità pari a 0.48 mm, prodotta a seguito dell'applicazione di un carico di compressione di 57 kN, teoricamente equivalente ad una pressione massima al centro del contatto pari a 2430 MPa.

A fianco si allegano le micrografie dello stato superficiale, scattate approssimativamente al centro del provino, prima e dopo l'esecuzione della prova di fatica.

Fratture in superficie

Si mostra, in Fig. 5, la composizione di alcune micrografie scattate sulla superficie del contatto dopo l'esecuzione di 2 milioni di cicli.

Le immagini evidenziano due anelli concentrici di cricche circonferenzialmente distribuite in corrispondenza del bordo dell'impronta di contatto, sulla superficie di un provino non lucidato.

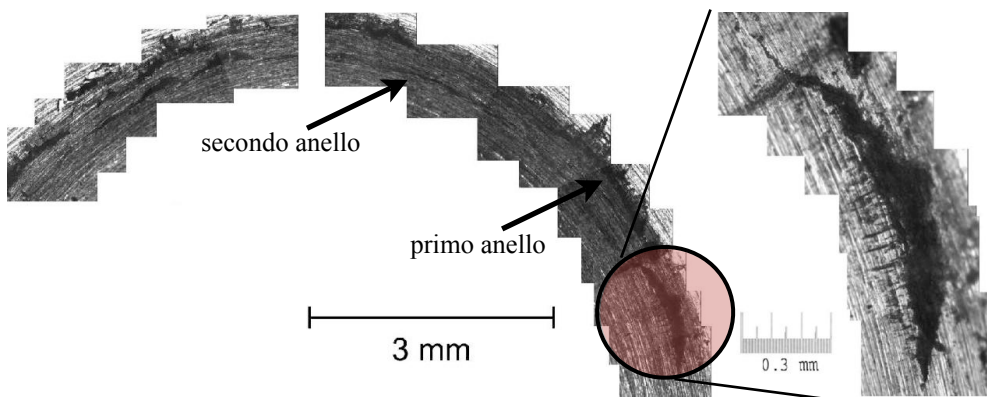


Figura 5. Cricche circonferenziali di superficie.

Si nota il confine tra il bordo dell'impronta (zona scura) e la parte piana della faccia del provino (zona più chiara), e si distinguono i solchi lasciati dall'utensile durante la tornitura della superficie. Le prime prove sperimentali sono state eseguite senza lucidare le facce piane delle provette cilindriche, in modo tale da sfruttare la rugosità superficiale per favorire l'innesco delle possibili fratture e migliorarne la visibilità. Confrontando gli esiti di tali prove esplorative con quelli delle prove effettuate sul materiale lucidato, che presentano la stessa morfologia di frattura, si è visto che l'effetto principale della rugosità superficiale consiste nell'accelerazione dell'innesco delle cricche e nell'aumento della velocità di accrescimento, senza tuttavia influenzarne la direzione di propagazione.

Sulla destra di Fig. 5 si riporta un ingrandimento di una zona di frattura spessa un decimo di millimetro circa, dal bordo della quale si vede partire una serie di cricche radiali minori. Si può motivare l'insorgenza di tali cricche "secondarie" a causa della pressione del carico di contatto, che spingendo il bordo ad occupare lo spazio libero lasciato dal volume della cricca "primaria", ne provoca la dilatazione e la conseguente frattura. Limitatamente alle osservazioni effettuate sui campioni sperimentali, si può affermare che questa tipologia di cricche radiali si presenta in via occasionale, come effetto secondario all'accrescimento delle fratture circonferenziali, rimanendo localizzata al bordo di queste ultime, senza dare luogo ad ulteriori propagazioni.

In Fig. 6, si mostrano le micrografie di una sezione dello stesso provino, scattate all'altezza della superficie, in corrispondenza del bordo dell'impronta plastica lasciata dall'indentatore sferico. Le frecce evidenziano i profili delle fratture di superficie circonferenzialmente distribuite secondo i due anelli concentrici descritti sopra.

Le fratture si propagano nella profondità del materiale, per un'estensione limitata a circa tre decimi di millimetro. Il fronte della cricca avanza incurvandosi in direzione dell'asse del contatto, e si arresta non appena viene raggiunta la profondità oltre la quale lo stato tensionale non risente più in modo significativo delle componenti di tensione di trazione, che sviluppano la loro massima intensità in superficie.

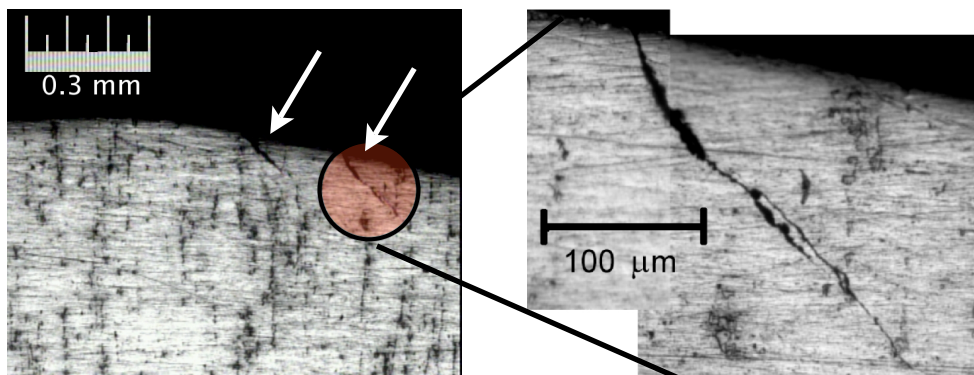


Figura 6. Micrografie della sezione di un provino.

Fratture sotto pelle

Le prove di fatica sul materiale duttile impiegato nella sperimentazione mostrano un ulteriore tipo di frattura, che si innesca sotto il bordo dell'impronta di contatto, ad una certa profondità dalla superficie.

In Fig. 7 si rappresenta uno schema qualitativo della sezione di una provetta, con la relativa traccia, non in scala, dell'indentazione plastica indotta dal contatto. Si evidenziano un punto, che rappresenta la zona di innesco della frattura, e due frecce, che indicano le direzioni di propagazione della stessa. La frattura si propaga sia verso il centro della provetta, sia verso il bordo laterale della stessa.

In Fig. 8 si riporta la composizione di una serie di micrografie scattate nella parte sinistra della sezione di una provetta, che mostra la morfologia di una frattura sotto pelle. Si evidenzia con la freccia il punto di innesco della frattura, che si propaga sia incurvandosi in direzione del centro, sia orizzontalmente in direzione della periferia. Nelle tre dimensioni, la frattura assume la geometria di rivoluzione, attorno all'asse del contatto, della traccia lasciata nella sezione di figura.

Gli aspetti critici di questo modo di guasto sono in primo luogo dovuti al fatto che la propagazione della cricca non si arresta, come avviene, invece, per le fratture di superficie (in assenza di lubrificazione). In secondo luogo, l'innesco e l'avanzamento del difetto sotto pelle sono tali da renderlo difficilmente rilevabile, se non impiegando metodi di indagine ispettivi quali, per esempio, la magnetoscopia, l'eco ad ultrasuoni o la radiografia, ove applicabili.

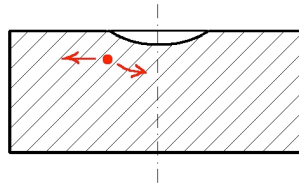


Figura 7. Schema di innesco e di propagazione della frattura sotto pelle.

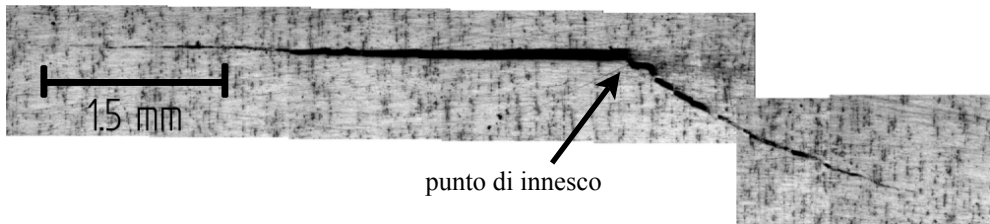


Figura 8. Micrografia di una sezione del provino. Frattura sotto pelle.

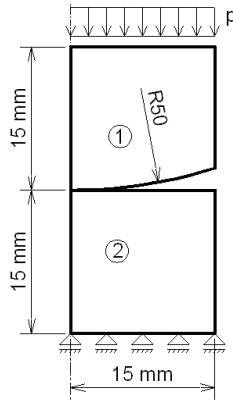


Figura 9. Modello FEM.

4. MODELLO NUMERICO DELLO STATO TENSIONALE NEL CONTATTO

Geometria

È stato realizzato un modello piano assialsimmetrico del contatto con il codice di calcolo ADINA, secondo la geometria di Fig. 9, dove il corpo nr. 1 rappresenta l'indentatore in acciaio ed il corpo nr. 2 la provetta di alluminio. L'asse del contatto coincide con il bordo sinistro del modello.

È stata presa in considerazione la condizione di carico relativa alle prove che hanno prodotto le micrografie discusse nelle sezioni precedenti, caratterizzata da una intensità massima del carico di 57 kN, corrispondente ad una distribuzione di pressione uniforme, come in figura, pari a 81 MPa.

Stato tensionale elastico

È stata effettuata una prima simulazione, considerando un modello lineare per le proprietà costitutive dei materiali, ed assumendo come coefficienti di proporzionalità i rispettivi moduli di elasticità. In queste condizioni si ritrovano, per lo stato tensionale, i risultati noti della teoria di Hertz.

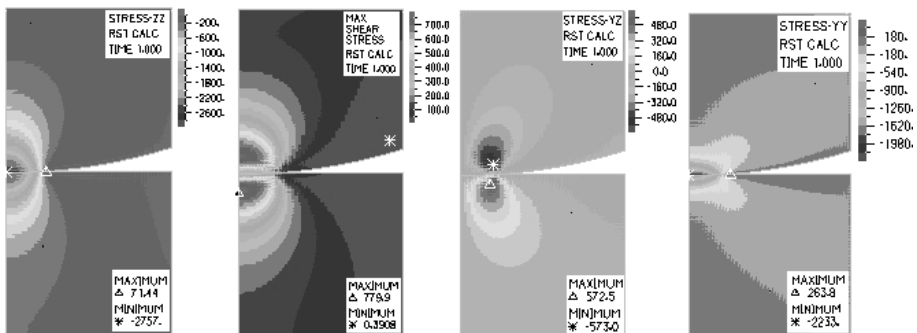


Figura 10. Stato tensionale elastico.

In Fig. 10, essendo y l'asse orizzontale e z l'asse verticale, si rappresentano, nell'ordine, gli andamenti della tensione normale nella direzione z del contatto, della massima tensione di taglio, della tensione di taglio nel piano yz di figura, e della tensione normale nella direzione y radiale.

Si osserva che la massima intensità della tensione tagliante è pari a 780 MPa, e che è localizzata ad una certa profondità, sull'asse del contatto, mentre la massima tensione di taglio nel piano yz è pari a 572 MPa, ed è situata sotto il bordo dell'impronta. Si vede, inoltre, che esiste una componente di tensione normale di trazione, che misura 264 MPa, disposta sulla superficie di contatto, al bordo dell'impronta, e diretta radialmente.

Secondo la teoria classica, la massima intensità del taglio è responsabile dei difetti che originano sotto pelle sull'asse z , mentre la massima tensione normale di trazione radiale è la responsabile dei difetti che si innescano in direzione circonferenziale, sulla superficie dei corpi, al bordo delle impronte di contatto.

Stato tensionale elasto-plastico

Per valutare numericamente lo stato tensionale tenendo conto delle deformazioni plastiche nel contatto, è stato effettuato un secondo studio, ipotizzando che il corpo in acciaio si mantenga in campo elastico, e simulando un legame costitutivo multi-lineare per il corpo in alluminio. Per quest'ultimo, sono state realizzate delle provette per prove di trazione quasi-statica, secondo la norma ASTM B577-94.

In Fig. 11 si mostrano la fotografia di una provetta rotta, con i lembi di frattura riaccostati, che evidenzia l'entità della strizione, e la curva di carico acquisita durante una prova. Nel diagramma sperimentale di figura, si mostra il legame fra la tensione, valutata dividendo il carico di prova per l'area della sezione nominale della provetta, e la deformazione del materiale, misurata mediante l'applicazione di un estensometro.

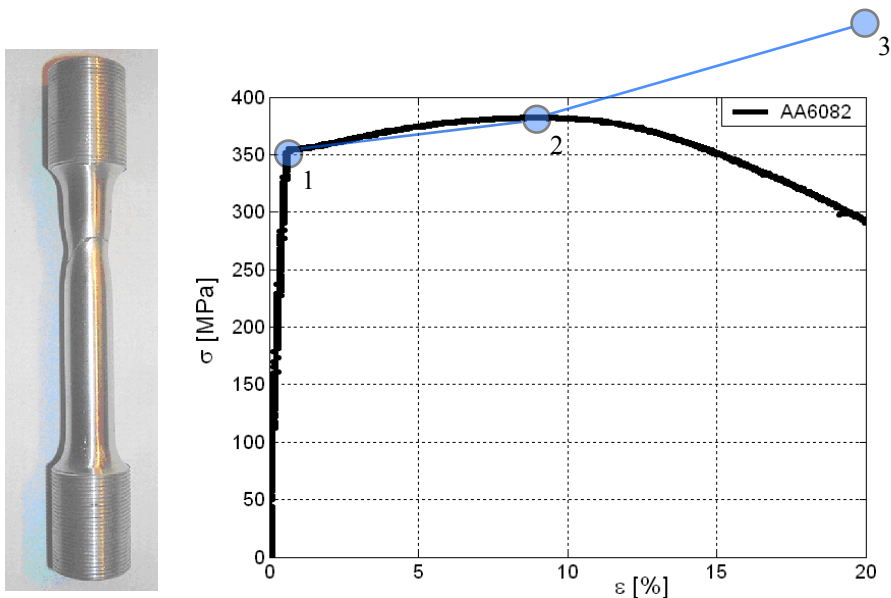


Figura 11. Provetta ASTM B577-94 e diagramma della prova di trazione.

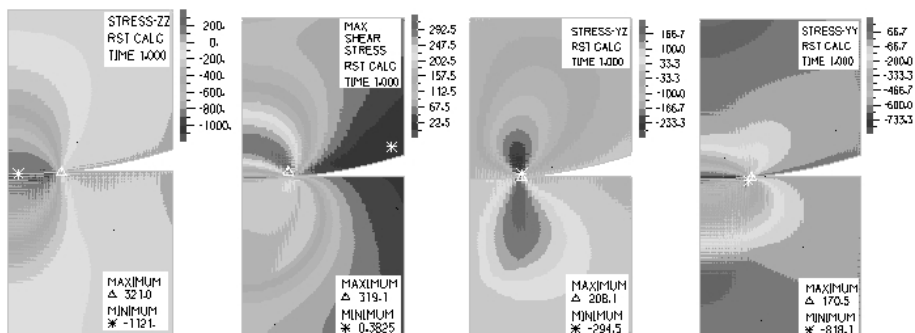


Figura 12. Stato tensionale elasto-plastico.

Ai sensi della simulazione numerica, tale diagramma è stato approssimato con l'andamento multi-lineare a tre pendenze sovrapposto in figura. Il primo tratto, dall'origine al punto 1, segue il comportamento del materiale in campo elastico; il secondo tratto, dal punto 1, dove il materiale inizia a snervarsi, al punto 2, nel quale inizia la strizione, descrive il comportamento tenso-deformativo del materiale in campo plastico. L'ultimo tratto si discosta dalla curva, in quanto viene tracciato dal punto 2 al punto 3 di rottura reale, nel quale la tensione ultima è stata calcolata dividendo il valore del carico di rottura per l'area della sezione minima di strizione. Tale valore misura 470 MPa.

Impiegando il modello costitutivo multi-lineare si ottiene il campo tensionale rappresentato in Fig. 12.

In figura, essendo y l'asse orizzontale e z l'asse verticale, si rappresentano ancora, nell'ordine, gli andamenti della tensione normale nella direzione z del contatto, della massima tensione di taglio, della tensione di taglio nel piano yz , e della tensione normale in direzione radiale.

Anche in questo caso si sviluppano tensioni radiali di trazione in corrispondenza della superficie del corpo, al bordo dell'impronta, mentre a differenza di quanto avviene per il caso della simulazione in campo elastico, la massima intensità della tensione di taglio non è localizzata in un punto sull'asse del contatto, bensì è diffusa in un ampio volume di materiale. Il suo valore misura circa 320 MPa, ed è confrontabile con il valore di picco della tensione tangenziale che si ha nel piano yz sotto il bordo dell'impronta, pari a 208 MPa, e che può quindi essere ritenuto responsabile delle fratture che si innescano in tale punto.

5. CONCLUSIONI

Sono state presentate e discusse alcune osservazioni sperimentali delle fratture che si innescano e propagano a seguito dell'applicazione di carichi affaticanti in un problema di meccanica del contatto in campo elasto-plastico. Tali fratture sono state messe a confronto con quanto citato in letteratura in merito al danneggiamento dei materiali in contatti in campo elastico.

Si è verificato che in entrambi i casi si manifestano rotture superficiali, distribuite lungo i bordi ellittici delle impronte. Tali fratture sono state messe in relazione con l'intensità delle componenti di tensione normale, che agiscono in direzione radiale, sulla superficie dei corpi, ed in corrispondenza dei suddetti bordi. È stata osservata la modalità

con la quale le cricche superficiali si propagano anche in profondità, rilevando come queste ultime ad un certo punto si arrestino, nel caso dell'applicazione di carichi affaticanti senza lubrificazione tra i corpi a contatto. Questo comportamento si discosta da quanto riportato in letteratura per il caso di contatti lubrificati, nei quali la pressione del fluido lubrificante che raggiunge l'apice della cricca può accelerarne la propagazione estendendola anche a profondità maggiori.

Si è osservato che l'innesco delle fratture sotto pelle si differenzia sensibilmente fra il caso dei "contatti elastici" e quello dei "contatti elasto-plastici", spostandosi, in questi ultimi, dall'asse del contatto al bordo dell'impronta. Questo comportamento è stato interpretato mediante uno studio agli elementi finiti, modellando il legame costitutivo del materiale tramite un andamento multi-lineare. La localizzazione del punto di innesco della frattura può così essere messa in relazione con la massima intensità della tensione di taglio, che si verifica nel piano della sezione verticale del corpo, sotto il bordo dell'impronta di contatto.

BIBLIOGRAFIA

- [1] Benedetti, A., Molari, P.G., Morelli, P., 2006. "Fracture Initiation and Propagation in Fatigue Contact Loading of AA6082 Aluminium Alloy". *Key Engineering Materials*, **324-325**, pp. 1091-1094.
- [2] Barber, J.R., Ciavarella, M., 2000. "Contact Mechanics". *International Journal of Solids and Structures*, **37**, pp. 29-43.
- [3] Shull, K.R., 2002. "Contact Mechanics and the adhesion of soft solids". *Material Science and Engineering*, **R36**, pp. 1-45.
- [4] Timoshenko, S., 1982. "Theory of elasticity".
- [5] Shah, K.R., Wong, T. "Fracturing at contact surfaces subjected to normal and tangential loads".
- [6] Panasyuk, V.V., Datsyshyn, O.P., Marchenko, H.P. 1995. "The crack propagation theory under rolling contact". *Engineering Fracture Mechanics*, **52**(1), pp. 179-191.

APPLICATION OF ADVANCED VIBRATION PROCESSING TECHNIQUES IN I.C. ENGINE COLD TESTS

Simone Delvecchio

*Engineering Department in Ferrara (EnDIF),
University of Ferrara, Italy
E-mail: simone.delvecchio@unife.it*

Gianluca D'Elia

*Engineering Department in Ferrara (EnDIF),
University of Ferrara, Italy
E-mail: gianluca.delia@unife.it*

Giorgio Dalpiaz

*Engineering Department in Ferrara (EnDIF),
University of Ferrara, Italy
E-mail: giorgio.dalpiaz@unife.it*

Abstract. *This work addresses the use of several signal processing tools as a means for the monitoring and the diagnosis of assembly faults in i.c. engines through the cold test technology. Firstly, an approach based on the use of symmetrized dot patterns for the visual characterization of vibration signatures is proposed. This method is applied in order to obtain reliable thresholds for the pass/fail decision after the cold test. Secondly, the fault identification is discussed on the basis of the cyclostationary modelling of the signals. First of all, first-order cyclostationarity is exploited through the analysis of the Time Synchronous Average (TSA). Subsequently, second-order cyclostationarity is developed by means of the mean instantaneous power, Wigner-Ville Distribution (WVD) and Wigner-Ville Spectrum (WVS). Finally, the Continuous Wavelet Transform (CWT) is presented and compared with the Wigner Ville Distribution. The effectiveness and the limitations of the above-mentioned techniques and diagnostic procedures applied to one considered assembly fault are compared and discussed on the basis of experimental results concerning a specific assembly fault.*

Keywords: *i.c. engines, vibration, monitoring, diagnostics, cyclostationarity, time-frequency distribution.*

1. INTRODUCTION

The purpose of condition monitoring is to use information extracted from the vibro-acoustic signature of a machine in order to detect faults or to define its state of health: a change in the vibration signature not only indicates a change in the machine conditions, but also directly points to the source of the signal alteration.

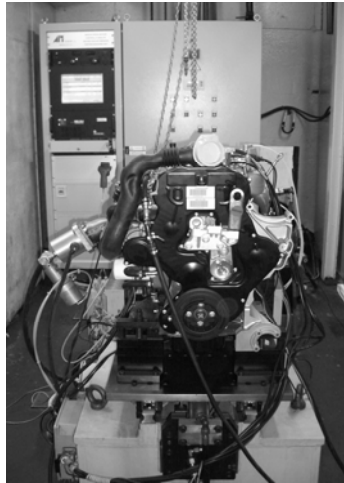


Figure 1. Cold test bench and data acquisition system.

Fault diagnosis, condition monitoring and fault detection are different terms sometimes improperly used. Condition monitoring and fault detection are always referred to the evaluation of the state of a machine and the detection of the anomaly. Fault diagnosis could be separated from the others because it is a more rigorous action which requires the identification of the component which causes the deviation from the normal state.

This paper deals with the vibration condition monitoring and fault diagnosis of assembly fault in diesel engine during cold tests. In this work only one faulty condition is discussed: it can be taken as a reference for a more general diagnostic procedure.

Nowadays, the main part of engine manufacturers test their engines by means of a “hot test”, i.e. a test in which the engine is firing. On the other hand, recently, some companies have introduced a “cold test” (see Figure 1) instead of the hot one, but this method has to be further improved.

The essential difference between these two methods is that the hot test aims to verify the engine performance, whilst the cold test aims to verify the anomalies by means of torque, pressure and vibration measurements.

It is worth noting that in the hot test only the main anomalies could be detected, while in the cold test the fault detection is more precise and effective, since no noise and vibration due to the firing are added.

In the “hot test”, each engine must be connected to a test bench which includes a brake and provides the cooling (water, oil) and the connection with the other auxiliary services in working condition as fuel, gas and air. For this reason, the cost in time for a complete hot test procedure is more than ten minutes. At present, the hot test technology only indicates which engines are not good for customers, but does not give any more information regarding the causes of the fault producing the malfunction.

Hot tests also give auxiliary costs for oil and fuel consumption, need of trained personal in order to set up the test, auxiliary costs for bench maintenance, and, as said above, difficulty for detecting the causes of the faults. On the other hand, the cold test reduces oil and combustible consumption, emissions of CO – SO₂ – CO₂ and test time until about 3 minutes.

As previously said, mechanical faults in machines often show their presence through abnormal vibration signals, thus techniques for machine condition monitoring based on the analysis of these signals are widely used [1-3].

However, most of the studies have been carried out on simple mechanical parts, such as gears and rolling bearings, having well determined dynamic characteristics. Therefore gear box condition monitoring and bearing defect analysis using vibration signatures are extensively reported (see Refs. [4, 5] among many works on these topics).

Vibration of reciprocating machinery such as internal combustion engines, compressors and pumps is far more complex and difficult to analyse. Moreover, the analysis of vibration signals of i.c. engines is affected by the complexity of the engine; some useful information can be deduced by Wowk [3], who asserts that “excessive vibration on reciprocating engines can be caused by operational problems such as misfiring, piston slap, compression leaks, faulty fuel injection and valve clash”. Therefore, each type of fault may produce a different vibration signature which can be detected by means of proper signal processing techniques.

Generally speaking, the fault detection and diagnosis of i.c. engines can be carried out using different strategies. One strategy can consist in modelling the whole mechanical system using lumped or finite element methods in order to simulate several faults and compare the results with what found through the experimental tests. Another strategy can adopt signal processing techniques in order to obtain features or maps that can be used to detect the presence of the defect. About this, a decision algorithm is needed for a visual or automatic detection procedure. Moreover the maps can be also analysed for diagnostic purposes. This method is the most commonly used and well suited for the judgment of expert technicians.

Most of the work in literature, pertaining with the above mentioned strategies, is referred to engines in firing condition. Kimmich et al. proposed a model-based diagnosis method [6]. Thomas et al. developed a diagnostic method to detect engine knock using pattern recognition with wavelet networks [7]. Molinaro et. al. used some pattern recognition detectors, such as cepstral coefficients and amplitude histograms, in order to improve knock recognition in spark engines [8]. Li et al. used the Independent Component Analysis to decompose noise signals into a number of independent components and, consequently, study each individual component separately [9].

Antoni et al. [10, 11] proposed a methodology based on the properties of cyclostationarity that was applied to malfunctions related to engines in firing condition (advance and delay of injection, misfires and knocks). In one of their work they use the vibration signal to derive the cylinder pressure overcoming the issue of fixing pressure transducers to the engine.

The application of time-frequency distribution techniques is well suited for the analysis of transitory signals and it has been widely applied to engine monitoring. The Short-Time Fourier Transform (STFT), the Wigner-Ville Distribution (WVD) and the Continuous Wavelet Transform (CWT) were used in order to distinguish faulty conditions from normal ones for practical fault diagnosis and not to obtain reliable parameters for an automatic procedure led by a data acquisition system [12-14].

This work tries to investigate the problem of obtaining a strategy for the pass/fail decision based on the comparison among different patterns and it seeks to localise the source of the alteration through the visual inspection of 2-D and 3-D maps.

The paper is organized as follows. Section 2 describes the experimental apparatus, the test conditions and the artificially introduced fault. Furthermore the characteristics of engine vibration are introduced. Section 3 concerns the application of the image correlation method

in order to find reliable thresholds for the pass-fail decision. Finally, Section 4 shows the cyclostationary model capabilities in localizing one specific assembly fault.

2. EXPERIMENTAL APPARATUS AND TEST CONDITIONS

Experimental investigations are carried out on a 2.8 L engine produced by VM Motori, 4-cylinder 4-stroke with four-valve-per-cylinder, turbocharged with an exhaust-driven turbo-compressor (firing order 1-3-4-2). As previously said, the measurements are made through the cold test technology in which the engine is driven by an electric motor via a coupling unit and therefore it is maintained in a non-combustion state.

The results presented in this paper deal with two different crankshaft speeds: the image correlation method was applied to signals extracted at 120 rpm while the cyclostationary and time-frequency analysis have been carried out on signals extracted at 1000 rpm. In order to obtain reliable comparisons all the vibration signals are picked up exactly at the same operating conditions. The vibration signal is measured by means of a piezoelectric general purpose accelerometer (MTN 1020; frequency range: 2-13 kHz) mounted on the engine block (turbocharger side) close to the bearing support of the crankshaft. The sample frequency was 14 kHz and the duration of the vibration signal acquisition was 2 s. The tachometer signal which measures the rotational position of the engine is generated using a crank sensor with 360 pulse/rev.

During the acquisition, the vibration signal is resampled with an angular resolution of 1 degree. Since the instantaneous speed has some fluctuations around the mean, the number of time samples varies from cycle to cycle, while the angular resampled signal is independent of the speed variations and can be used to compute the Time Synchronous Average (TSA) over 16 time histories, each corresponding to two crankshaft rotations (720°), in order to respect the periodicity of a 4-stroke engine.

This synchronous average completely depicts the nature of the signal. As shown by Antoni [10] this cyclic average allows the application of cyclostationarity tools such as the WVD. The signal measured in healthy conditions can be assumed as the superposition of periodic components due to crankshaft rotation, piston motion, valve activities, and a transient part due to the pressurization of the cylinders.

In order to evaluate the vibration signal diagnostic capabilities tests were carried out with different types of fault, applying a number of signal processing techniques.

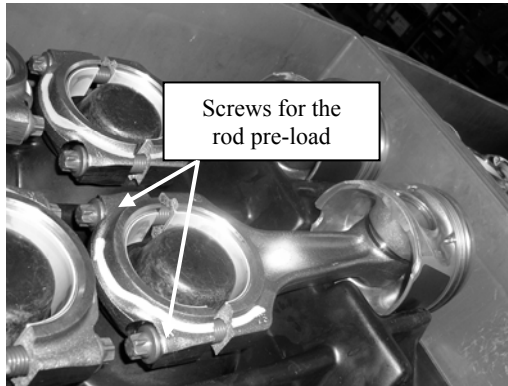


Figure 2. Rod screws incorrectly preloaded.

In this paper the effectiveness of these techniques are discussed and compared with reference to a specific fault, as example:

- Rod with screw pre-load of only 3 kgm, instead of the correct torque of 9 kgm (RS3 engine) – Figure 2. This fault causes an irregular rotation of the rod and, consequently, an incorrect stroke of the piston, this affecting the engine operations. The faulty rod is connected to the third piston in the firing order.

3. CONDITION MONITORING USING SYMMETRIZED DOT POLAR GRAPHS AND IMAGE CORRELATION

Signal processing

As presented in Ref. [15], the normalized time histories of the vibration signal concerning diesel engines can be represented as symmetrized dot graphs.

The formulation used for the transformation of the discrete signal $x(n)$ to a polar coordinate graphs and the properties of this method are described in [16]. In this work the time lag L and the gain ζ present the following values: $L=1$, $\zeta =60^\circ$. Examples of the results are shown in Figure 3 and Figure 4. In order to implement this technique in the cold test procedure for fault detection, it is necessary to develop an image correlation system. The authors applied the algorithm of the edge detection, illustrated in [17], which represents the most common approach for detecting meaningful discontinuities in intensity values. The basic idea behind the edge detection is to find the points where the intensity rapidly changes.

For each case, by applying this edge detection algorithm on the image of the visual symmetrized dot polar graph, the result will be a logic matrix with 1's (represented as white pixels in the grey scale) at the locations where edge points were detected on the image, and 0's (black pixels) elsewhere (Figure 3 (c)).

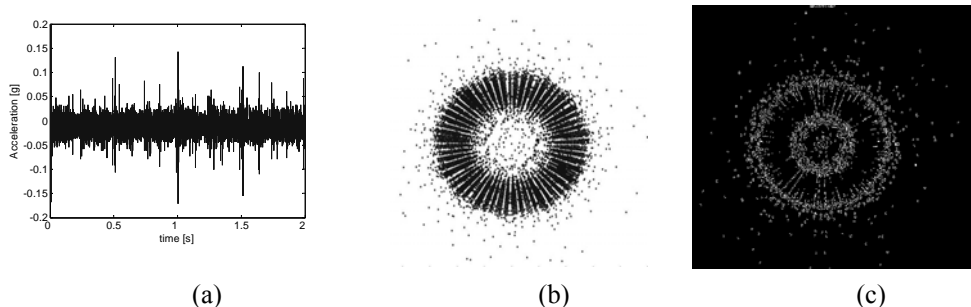


Figure 3. Normal pattern: (a) time input waveform; (b) symmetrized dot polar graph; (c) image obtained after the application of the edge detection algorithm.

The goal is to identify a reference pattern that represents the normal condition and then compare the images obtained from all the test engines with this ‘healthy pattern’ by means of a similarity parameter. Among many possibilities, this parameter, called correlation, was chosen as the percentage of common white pixels with respect to the total number of white pixels in the healthy engine pattern.

Experimental results

In order to set the detection threshold for the correlation, first of all this parameter is calculated for all the couples of engines among a group of 21 healthy engines: the engine that presents the minimum correlation with each other is assumed as the reference pattern, unique for all investigations, and the corresponding minimum correlation of 25.08% is fixed as the correlation threshold (CT).

This pattern is compared with the images obtained from different faulty engines and so the correlations are calculated, verifying if they are lower than the CT in order to discriminate the faulty condition from the normal one. Hereafter only the result concerning the faulty engine with incorrectly preloaded rod screws is shown, as example: Table 1 shows that this faulty engine gives a percentage of common white pixels lower than the CT. Similar results were found for other faults. Thus, it seems that this technique exhibits high sensitivity to the fault detection presenting some advantages: fast computational implementation and easy output evaluation.

4. FAULT DIAGNOSTICS USING CYCLOSTATIONARY ANALYSIS

Formulation

The engine time vibration signal $x(t)$ can be considered as a wide sense cyclostationary process that requires the periodicity of the first two moments.

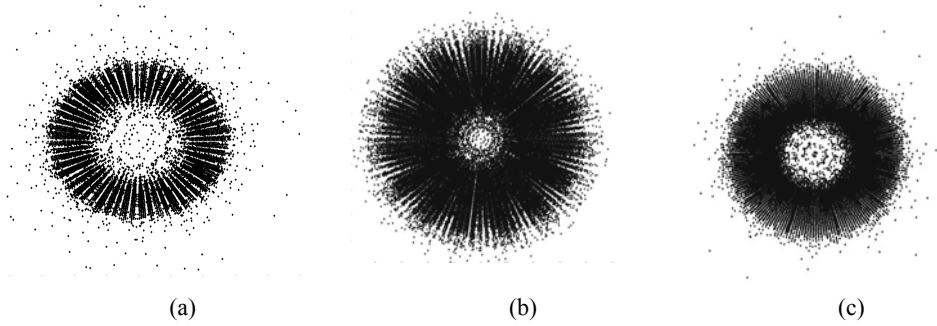


Figure 4. Symmetrized dot polar graphs: (a) healthy engine (reference pattern); (b) and (c) faulty engines.

Table 1. Symmetrized patterns: correlation values.

Comparison between symmetrized dot polar patterns	Percentage of common white pixels [%]
Healthy engine - Healthy engine	25.08 (CT)
Healthy engine - Faulty engine (incorrectly preloaded rod screws)	10.80

First order cyclostationarity

$$m_x(t) = E\{x(t)\} = m_x(t+T) \quad (1)$$

Second order cyclostationarity

$$R_{xx}(t_1, t_2) = E\{x(t_1)x^*(t_2)\} = R_{xx}(t_1+T, t_2+T) \quad (2)$$

where $E\{\cdot\}$ stands for the expected value and T is the time period.

The Wigner Ville Distribution (WVD) of the time signal is defined as follows:

$$\begin{aligned} W_x(t, f) &= \int_{-\infty}^{+\infty} x\left(t + \frac{\tau}{2}\right) \cdot x\left(t - \frac{\tau}{2}\right)^* \cdot e^{-j2\pi\tau f} d\tau = \\ &= FT_{\tau \rightarrow f} \left\{ x\left(t + \frac{\tau}{2}\right) \cdot x\left(t - \frac{\tau}{2}\right)^* \right\} \end{aligned} \quad (3)$$

with FT refers to the Fourier Transform.

The Wigner Ville Spectrum (WVS) is the expected value of the WVD:

$$\begin{aligned} WVS_x(t, f) &= E \left\{ FT_{\tau \rightarrow f} \left\{ x \left(t + \frac{\tau}{2} \right) \cdot x \left(t - \frac{\tau}{2} \right)^* \right\} \right\} = \\ &= FT_{\tau \rightarrow f} \left\{ E \left\{ x \left(t + \frac{\tau}{2} \right) \cdot x \left(t - \frac{\tau}{2} \right)^* \right\} \right\} \end{aligned} \quad (4)$$

As can be noted the WVS is the Fourier Transform of the Instantaneous Auto-Correlation function:

$$R_x(t, f) = E \left\{ x \left(t + \frac{\tau}{2} \right) \cdot x \left(t - \frac{\tau}{2} \right)^* \right\} \quad (5)$$

Antoni [10] interprets the WVS (under the assumption of non-negativity) as the probability density function of the frequency random variable ν vs. the angular position variable (see Figure 7). Because of this, the WVS can be described by its “spectral” moments.

The n th spectral moment is defined as:

$$m_n^x(t) = \int_0^{\infty} WVS_x(t, \nu) \nu^n d\nu \quad (6)$$

Antoni proves that the spectral moments can be obtained from the following formulation

$$m_n^x(t) = (2\pi j)^{-n} \frac{\partial^n \ln K}{(\partial \tau)^n}(t, 0) \quad (7)$$

where K is the autocovariance function of the analyzed process.

The estimated sampled mean instantaneous power (MIP) can be obtained directly from the angular sampled signal:

$$\hat{m}_0^x[m] = \left\langle |X_a[m]|^2 \right\rangle_N^I \quad m = 0, \dots, N - 1 \quad (8)$$

with X_a the analytic signal where $\left\langle |\cdot| \right\rangle_N^I$ means averaging over I cycles of length N samples of a single measurement.

Application to a signal model

In order to make quantitative evaluations about the cyclostationary analysis capabilities, a signal model, that realistically mimics the vibration signal measured from the cold test, is firstly analytically formulated:

$$x(t) = x_d(t) + x_{nd}(t) + x_s(t) \quad (9)$$

where $x_d(t)$ and $x_{nd}(t)$ are respectively the deterministic and the non-deterministic part. The non-deterministic part $x_{nd}(t)$ considers the random fluctuations around $x_d(t)$, while $x_s(t)$ takes into account the background noise.

Using this model proposed by Antoni [10] we want to asses the diagnostic procedure for measurements carried out in a non-combustion state.

The deterministic part $x_d(t)$ is first-order cyclostationary:

$$x_d(t) = x_d(t + T) \quad (10)$$

and $x_{nd}(t)$ is set to be second-order cyclostationary respecting the following condition expressed in terms of autocovariance function $K_{x_{nd}x_{nd}}$:

$$K_{x_{nd}x_{nd}}(t, \tau) = K_{x_{nd}x_{nd}}(t + T, \tau) \quad (11)$$

where T is the non-random period.

The autocovariance function is expressed by:

$$K_{x_{nd}x_{nd}}(t, \tau) = E \{ x_{nd}(t + \tau/2) x_{nd}(t - \tau/2) \} - \mu_{nd}(t + \tau/2) \mu_{nd}(t - \tau/2) \quad (12)$$

where μ_{nd} is the mean of the non deterministic part $x_{nd}(t)$.

The deterministic part can be modelled as a series of impulse response functions such as:

$$x_d(t) = \sum_{k=1}^N x_{dk}(t) \quad (13)$$

$$\text{with } \begin{cases} t < t_{impk} \rightarrow x_{dk}(t) = 0 \\ t \geq t_{impk} \rightarrow x_{dk}(t) = e^{-(t-t_{impk})/\sigma_k} X_{dk} \cos[2\pi f_{impk}(t-t_{impk})] \end{cases}$$

where X_{dk} and t_{impk} define respectively the amplitude and the time of the k^{th} engine event, whilst σ_k is the time constant related to the damping factor of the k^{th} impulse response.

The non-deterministic signal counterpart is modelled as a cyclic impulse modulated noise, such as:

$$x_{nd}(t) = \sum_{k=1}^N x_{ndk}(t) \quad (14)$$

$$\text{with } \begin{cases} t < t_{impk} \rightarrow x_{ndk}(t) = 0 \\ t_{impk} \leq t < t_{imp(k+1)} \rightarrow x_{ndk}(t) = X_{ndk} e^{-(t-t_{impk})/\sigma_k} n(t) \end{cases}$$

where $n(t)$ is a normally distributed random noise of zero mean and standard deviation one, whilst t_{impk} and σ_k are defined as above; $x_s(t)$ is still modelled as a normally distributed random noise with zero mean and standard deviation one.

The specific values of the parameters used in the signal model synthesis are listed in Table 2 and Table 3.

We want to remark that the physical periodicity really belongs to the crankshaft angle and not to time. Therefore the angular signal is well suited for the cyclostationary modelling. So, the vibration signal model of Equation (9) can be expressed as follows:

$$x(\theta) = x_d(\theta) + x_{nd}(\theta) + x_s(\theta) \quad (15)$$

where θ is the crankshaft angular position. Following this model, the WVD and WVS will be applied in the angle domain, although the above-reported formulation has been treated in the time domain. Moreover both representations will be plotted scaling frequency in Hz rather than in deg-1. This fact makes sense because a constant engine speed is assumed (1000 rpm).

Due to the fact that the measurements are carried out in a non-combustion state, the vibration signal model only considers the engine pressurization and inertial forces in healthy conditions. In order to simulate a faulty condition, an impulsive component is added to the non-stationary signal counterpart and simulated as an impulse response (see Table 3, k=2). The result is shown in Figure 5.

The signal model (see Figure 6) was considered satisfactory to point out the different contributions of the deterministic and cyclostationary parts on the total signal. Moreover the computation of the mean instantaneous power of the residual signal seems to be effective in highlighting the presence and the angular location of a generic fault.

Table 2. Vibration signal model – Deterministic part.

k	1	2	3	4
X_{dk}	2	2	2	2
σ_k [s]	0.0017	0.0017	0.0017	0.0017
t_{impk} [s]	0	0.031	0.062	0.093
f_{impk} [Hz]	4000	4000	4000	4000

Table 3. Vibration signal model – Cyclostationary part.

k	1	2	3	4	5
X_{ndk}	0.5	2	0.5	0.5	0.5
σ_k [s]	0.0017	0.0017	0.0017	0.0017	0.0017
t_{impk} [s]	0	0.0155	0.031	0.062	0.093

Experimental results

In order to carry out the cyclostationarity analysis, the time signal is synchronously resampled and the time synchronous average (TSA) is computed over two crankshaft rotations (Figure 8 (a)). In this case, the comparison among healthy and faulty RMS and Peak values (Table 4) does not reveal the presence of any defect. As it happens both faulty RMS and Peak signal values are below the healthy ones. Nevertheless, the acceleration peaks of the synchronous average can be physically explained by inspecting the engine event diagram.

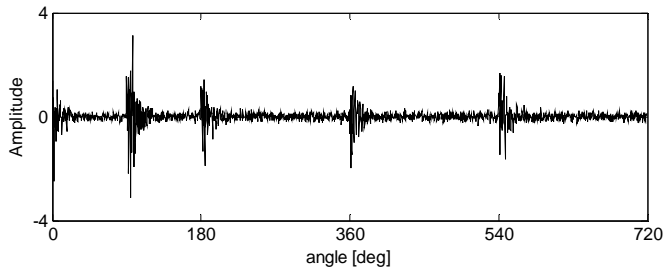


Figure 5. Model of i.c. engine vibration signal, considering only pressurization, inertial forces and an impulsive component.

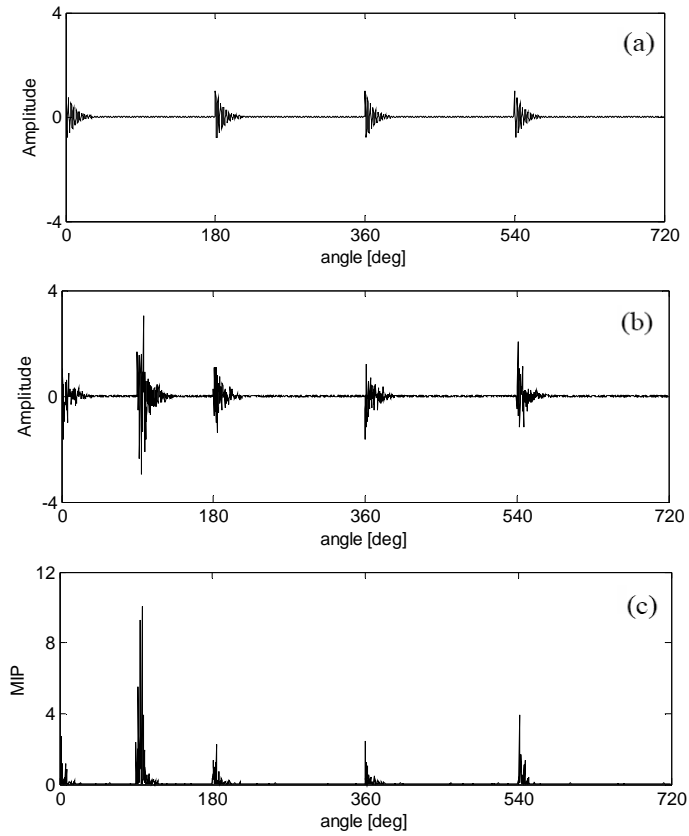


Figure 6: Model of i.c. engine vibration signal, considering only pressurization, inertial forces and an impulsive component: (a) deterministic part; (b) cyclostationary part; (c) mean instantaneous power of the residual signal.

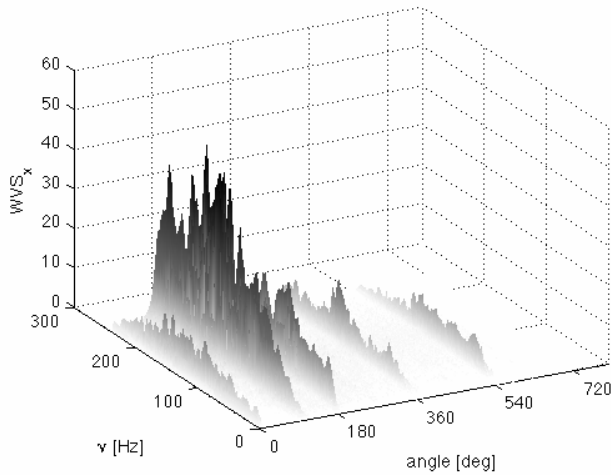


Figure 7. WVS of the cyclostationary part, represented in Figure 6(b).

As shown in Figure 8(a) acceleration peaks are present on the pressurization of each cylinder.

Moreover other acceleration peaks appear at the closure of the intake valves. In fact, due to the incorrectly tight rod, there are wide clearances between rod and crankshaft. These clearances are abruptly traversed whenever there is a change in the verse of the resultant force acting on the rod. In this case the resultant force takes into account the contribution of the pressurization and the inertial force.

The angular mismatch in the acceleration peaks and the closure of the intake valves of cylinders 3 and 4 is due to the ordinary delay on the closure of the intake valves which takes place in practice. As a result, this type of analysis is inadequate to predict the presence and position of the fault. For that reason, a second order analysis is performed.

The mean instantaneous power (MIP) of the residual signal is evaluated and plotted in Figure 8(c), which also displays the faulty Peak value compared with the healthy one. As one can clearly see, the presence of the defect is confirmed by the differences in the MIP Peak values: the MIC Peak value is approximately five times the healthy one.

Moreover, by the comparison with the engine event diagram one can notice that the largest energy amount is located in correspondence of the closure of the third cylinder intake valves, which points out the defect location. We want to remark that the closure of the valve is taken as an angular reference in order to understand the angular location of the crankshaft with respect to the firing order. At this stage, it can be concluded that the analysis of the mean instantaneous power makes know that the incorrectly tight rod is located in the third cylinder.

In order to confirm these results, the WVD of the synchronous average and the residual signal are computed. Figure 9(a) and (b) plot the WVDs of the synchronous average and residual signal respectively.

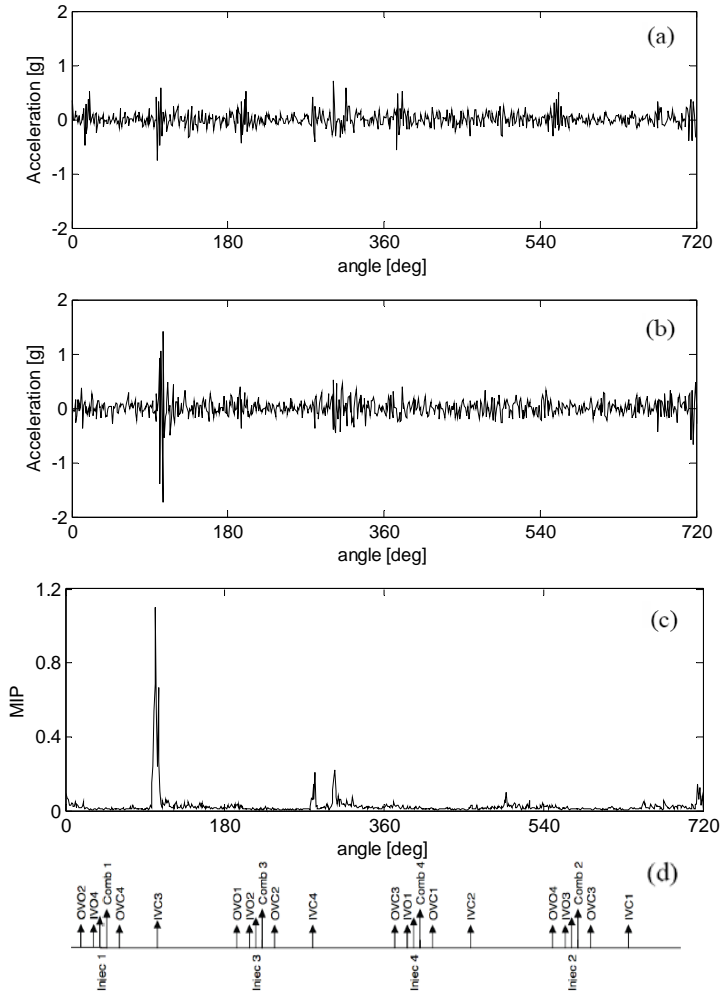


Figure 8. Incorrectly preloaded rod screws: (a) time synchronous average, (b) residual signal, (c) mean instantaneous power of the residual signal, (d) engine event diagram.

The WVD of the synchronous average is of difficult interpretation. De facto vertical lines appear at the pressurization of each cylinder and at the closure of the intake valves. Unfortunately, because of the acceleration amplitude associated to the engine events are about two times the back-ground noise (see Figure 8(a)), the vertical lines associated to the engine events are not clearly visible in the WVD, and in addition several cross-terms of relevant amplitude arise in the distribution. Therefore, the fault presence and its location cannot be established from Figure 9(a).

For that reason the WVD of the residual signal is computed (see Figure 9(b)). As previously said, the residual signal takes into account the 2nd order cyclostationary content of the signal, and so, by computing the WVD of this signal, one can obtain a distribution in which the energy content is mainly restricted at the fault location.

Table 4. Statistics extracted from TSA and MIP.

	RMS		Peak	
	Normal	Faulty	Normal	Faulty
Vibration Time Synchronous Average [g]	0.17	0.14	1.20	0.71
Mean Instantaneous Power (MIP) [g ²]	-	-	0.22	1.10

In fact, Figure 9(b) has only a vertical line in correspondence of the closure of the intake valves of the third cylinder.

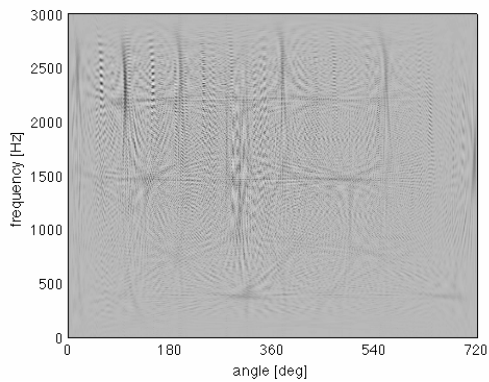
The CWTs (impulse mother wavelet is used) of both synchronous average and residual signal are also computed (Figure 9 (c) and (d)). It can be noted that the CWT of the time synchronous average is clearer than the WVD of the same signal. In particular the CWT map of the synchronous average (Figure 9(b)) detects the four cylinder pressurizations and the third cylinder intake valve closure. Even if a remarkable vertical line corresponding to the IVC3 is present in the CWT map of the synchronous average (Figure 9(b)), it is not enough in order to assure the presence of a mechanical fault. Therefore, the CWT of the residual signal is an expected step for the mechanical fault localization. As depicted in Figure 9(d) the presence of the pre-loaded rod is highlighted by the marked vertical line at about 100°. Hence, as previously observed in WVD analysis, the fault location can be only achieved by the analysis of the residual signal. It is worth noting that the WVD of the residual signal, taking the advantages of the best time-frequency resolution everywhere, show a better time fault localization compared to the ones obtained by the CWT.

Finally, the last proof of the mechanical fault presence can be obtained by the WVS of the residual signal. De facto, Figure 10 shows only a vertical line corresponding to the IVC3 highlighting so the fault location. The WVS, joining the advantages of the best time-frequency resolution and the cyclic averaging power of the autocorrelation function is also able to point out other engine events happening within the entire engine cycle. This can be considered the best representation of the time-frequency response within an engine cycle.

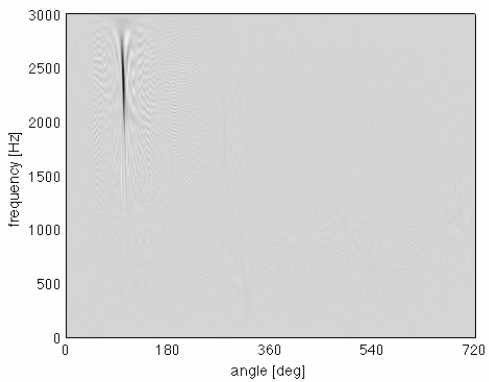
5. CONCLUDING REMARKS

This paper addresses the use of some advanced vibration processing techniques in analyzing acceleration signals measured from diesel engine cold tests for fault diagnostics purposes. Tests were carried out for a number of different faults. In this paper, the results relative to a specific fault are shown, as example. The fault concerns the incorrectly preloaded rod screws and causes incorrect engine operations and hits.

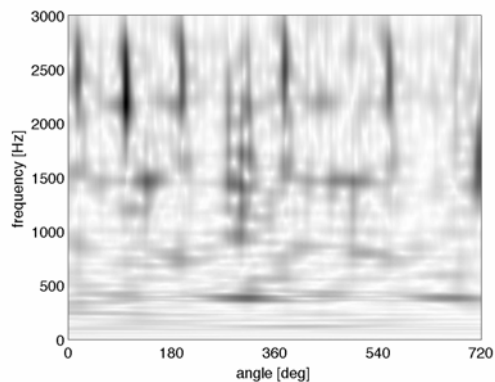
Firstly, the experimental apparatus concerning the vibration tests is explained, highlighting the benefits and the drawbacks of the cold test compared to the genuine hot test. After that the vibration signal model for this type of faults is introduced. It deals with cyclostationary model proposed by Antoni et al. in [10], in which the signal is subdivided into two main parts: deterministic and non-deterministic. Finally, the acceleration signals acquired from the engine block during a cold test cycle at the end of the assembly line are analyzed. For quality control purposes, an image matching correlation of symmetrized dot graphs is described. This method visualises vibration signals in a diagrammatic representation in order to quickly detect the faulty engines in cold tests.



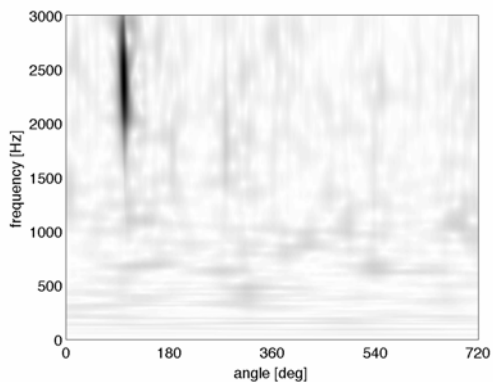
(a)



(b)



(c)



(d)

Figure 9. Incorrectly preloaded rod screws: (a) WVD of the TSA; (b) WVD of the residual signal; (c) CWT of the TSA; (d) CWT of the residual signal.

The percentage of common white pixels with respect to the total number of white pixels in a healthy pattern, taken as a reference, can be considered as a reliable option for monitoring purposes.

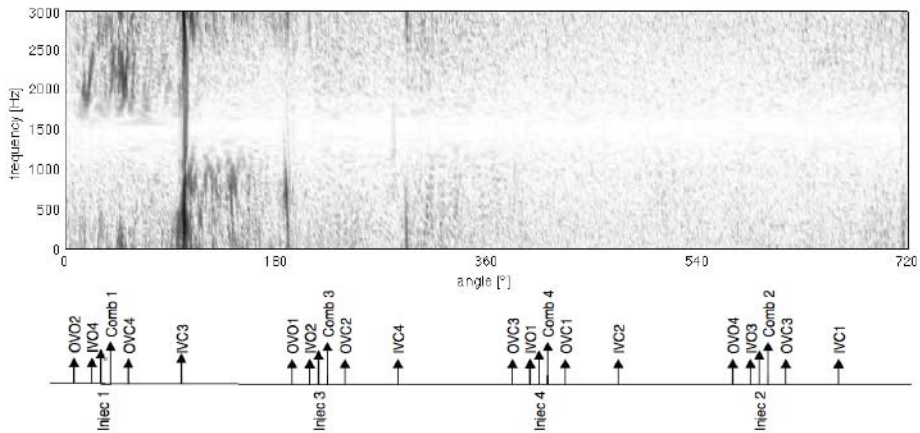


Figure 10. Incorrectly preloaded rod screws: WVS of the residual signal.

For the case under observation, the signal is synchronously sampled and the time synchronous average is computed over two crankshaft rotations. Subsequently the residual signal is evaluated by subtracting, from the synchronized signal, the time-synchronous average. In fact the periodicities due to the normal engine operating conditions are zeroed in

the residual signal, which can therefore point out the power linked to the fault. The residual signal, when necessary, is also analyzed by means of the WVS and the mean instantaneous power, as suggested by Antoni et al. in [10].

For the fault considered in this paper, the analysis of the residual signal is needed. The analysis carried out on the residual signals shows the presence and the location of the assembly fault. In particular the mean instantaneous power, evaluated over the residual signal, is strongly influenced by the fault location. Therefore, this technique can be a powerful tool in vibration based diagnostics. In fact, the mean instantaneous power is of simple numerical implementation, so interesting results can be obtained without the use of other time frequency techniques of more difficult implementation. However, in order to confirm the detection ability of the mean instantaneous power, the WVD and WVS are here also computed. The vertical lines associated to the engine events are not clearly visible in the WVD of the synchronous resampled signal; in addition several cross-terms of relevant amplitude arise in the distribution. Therefore, it could be concluded that the time-frequency analysis of the time-synchronized signal carried out by means of WVD is not suitable for the detection of the fault location, because this signal represents the deterministic signal content, which is strictly linked to the engine events. As a result the analysis of the residual signal is of pivotal importance in vibration based diagnostics. The CWT of the TSA was already able to indicate clearly the fault location. The CWT of the residual signal has shown a more spread angular fault location. Concerning the WVS computation, it has shown the highest sensitivity to fault detection. In fact, as depicted in Figure 10, the assembly fault location can be strongly highlighted by the appearance of vertical lines at the angular position of the rod affected by that fault. In addition, this technique is also able to show the main engine events.

Acknowledgements

The authors wish to thank Apicom and VM Motori (Cento, Ferrara, Italy) and the engineers of these Companies for co-operation and assistance in the collection of engine data. This work has been developed within the laboratory of research and technology transfer InterMech (Division Acoustics and Vibration - LAV) realized through the contribution of Regione Emilia Romagna - Assessorato Attivita' Produttive, Sviluppo Economico, Piano telematico, PRRIITT misura 3.4 azione A.

REFERENCES

- [1] Braun, S. J., 1986. *Mechanical Signature Analysis*. Academic Press, London.
- [2] Collacot, R.A., 1979. *Vibration Monitoring and Diagnostic*. Wiley, New York.
- [3] Wowk, V., 1991. *Machinery vibration: measurement and analysis*. McGraw-Hill, New York.
- [4] McFadden, P.D., 1987. "Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration". *Mechanical Systems and Signal Processing*, **1**(2), April, pp. 173–183.
- [5] Randall, R.B., 1982. "A new method of modelling gear faults". *Journal of Mechanical Design*, **104**, pp. 259–267.
- [6] Kimmich, F., Schwarte, A., and Isermann, R., 2005. "Fault detection for modern Diesel engines using signal and process model-based methods". *Control Engineering Practice*, **13**, pp. 189-203.
- [7] Thomas, J., Dubuisson, B., and Dillies-Peltier, M.A., 1997. "Engine Knock Detection from Vibration Signals using Pattern Recognition". *Meccanica*, **32**, pp. 431-439.
- [8] Molinaro, F., and Castanie, F., 1995. "Signal processing pattern classification techniques to improve knock detection in spark ignition engines", *Mechanical Systems and Signal Processing*, **9**(1), pp. 51-62.
- [9] Li, W., Gu, F., Ball, A. D., Leung, A. Y., and Phipps, C. E., 2001. "A study of the noise from diesel engines using the Independent Component Analysis". *Mechanical Systems and Signal Processing*, **15**(6) pp. 1165-1184.
- [10] Antoni, J., Daniere, J., and Guillet, G., 2002. "Effective vibration analysis of ic engines using cyclostationarity. Part I-A methodology for condition monitoring". *Journal of Sound and Vibration*, **257**(5), pp. 815-837.
- [11] Antoni, J., Daniere, J., and Guillet, G., 2002. "Effective vibration analysis of ic engines using cyclostationarity. Part II-New results on the reconstruction of the cylinder pressures". *Journal of Sound and Vibration*, **257**(5), pp. 839-856.
- [12] Mahjoob, M. J., Zamanian, A., 2006. "Vibration Signature Analysis for Engine Condition Monitoring and Diagnosis". In Proceedings of ISMA2006 International Conference on Noise and Vibration Engineering, P. Sas, B. Bergen, eds., Leuven, Belgium, pp. 885-896, Paper number 604.
- [13] Geng, Z., Chen, J., and Barry Hull, J., 2003. "Analysis of engine vibration and design of an applicable diagnosing approach". *International Journal of Mechanical Sciences*, **45**, pp. 1391-1410.

- [14] Tse, P.W., Yang, W., and Tam, H.Y., 2004. "Machine fault through an effective exact wavelet analysis". *Journal of Sound and Vibration*, **277**, pp. 1005-1024.
- [15] Wu, J. D., and Chuang, C.Q., 2005. "Fault diagnosis of internal combustion engines using visual dot patterns of acoustic and vibration signals". *NDT&E International*, **38**, pp. 605-614.
- [16] Shibata, K., Takahashi, A., and Shirai, T. 2000. "Fault diagnosis of rotating machinery through visualisation of sound signal". *Mechanical Systems and Signal Processing*, **14**(2), pp. 229-241.
- [17] In Gonzalez, R., Woods, R., and Eddins, S., 2004. *Digital Image Processing Using Matlab*. Prentice Hall.

APPLICAZIONE DELL'OSCILLATORE DI DUFFING PER IL MONITORAGGIO DELLE MACCHINE

Umberto Meneghetti

DIEM – Dipartimento di Ingegneria delle costruzioni meccaniche, nucleari, aeronautiche e di Metallurgia - Università di Bologna

e-mail: umberto.meneghetti@unibo.it

Alberto Maggiore

DIEM – Dipartimento di Ingegneria delle costruzioni meccaniche, nucleari, aeronautiche e di Metallurgia - Università di Bologna

e-mail: alberto.maggiore@unibo.it

Sommario. *Fra le tecniche utilizzate per il monitoraggio dello stato di salute delle macchine, l'analisi del segnale vibratorio è senz'altro quella più comunemente impiegata. Un problema spesso presente è quello relativo alla necessità di rilevare un segnale debole immerso in un rumore di elevata intensità, come, ad esempio, il segnale vibratorio dovuto alla presenza di un difetto allo stadio iniziale, al quale, in un ambiente reale, si sovrappongono di regola altri segnali di diversa natura e provenienza. Recentemente è stato proposto l'impiego di un sistema caotico, il cui comportamento è altamente sensibile alla presenza di un segnale sinusoidale di opportuna frequenza, anche di piccola intensità. Nel presente lavoro viene richiamata la metodologia suddetta e viene proposto un nuovo metodo per la valutazione numerica del caos. L'applicazione ad alcuni casi, con segnali sia virtuali sia sperimentali, e il confronto con altre metodologie, permettono una prima valutazione delle possibilità e dei limiti della metodologia proposta.*

Parole chiave: Duffing, caos, monitoraggio, diagnostica, segnale

1. INTRODUZIONE

Il monitoraggio e la diagnostica delle macchine mediante il rilievo e l'analisi delle vibrazioni sono entrati da tempo nell'uso corrente e si stanno diffondendo anche per il controllo della qualità.

Per questi scopi sono comunemente impiegate numerose tecniche di analisi del segnale, dalle più semplici alle più complesse e sofisticate; la descrizione di tali tecniche esula dagli interessi di questo lavoro: per essa rimandiamo ai testi e alle riviste citati in bibliografia, v. per es. [1 – 6].

È naturalmente desiderabile poter rilevare la presenza di un danneggiamento o di un difetto anche quando il relativo segnale è debole, come per es. nel caso di un danno allo stato iniziale, o quando il segnale utile è mascherato da un rumore elevato, come capita frequentemente negli ambienti industriali.

Ciò giustifica la ricerca di tecniche atte a rilevare in un segnale la presenza di una determinata componente, anche quando questa sia di debole intensità: di solito, si tratta di un segnale periodico di frequenza nota, riferibile al danno o al difetto che si desiderano rilevare.

Recentemente è stato proposto a questo fine l'impiego di sistemi soggetti a comportamento caotico, in particolare dell'oscillatore di Duffing [7 - 13]: sfruttando l'elevata sensibilità di questi sistemi alle condizioni iniziali [14], si può infatti rilevare la presenza di una debole componente periodica, per es. sinusoidale, anche in un segnale molto rumoroso.

La presenza e – in una qualche misura – l'entità di tale componente sono rivelate dal passaggio del sistema dallo stato di caos a quello di non caos, o viceversa: è perciò importante poter individuare l'eventuale presenza del comportamento caotico mediante un parametro quantitativo, cioè mediante un numero.

Nel presente lavoro, dopo alcuni brevi richiami sull'oscillatore di Duffing, sono descritte per sommi capi alcune metodologie del suo impiego ai fini del monitoraggio e della diagnostica industriale; si propone quindi un nuovo metodo per quantificare il caos e si descrivono i relativi risultati su segnali artificiali; si passa infine all'applicazione ad alcuni casi reali e si commentano i risultati ottenuti.

2. L'OSCILLATORE DI DUFFING

Come è ben noto [15], per studiare il comportamento oscillatorio dei sistemi non lineari si impiega comunemente il piano delle fasi: data l'importanza di tale rappresentazione per il seguito della presente trattazione, riteniamo opportuno richiamare innanzi tutto il caso dell'oscillatore lineare.

Consideriamo pertanto un sistema lineare non smorzato massa-molla; posto per comodità uguale a uno il rapporto fra la massa e la costante elastica della molla, senza con ciò perdere di generalità, e indicato con x lo spostamento della massa, l'equazione del moto libero è:

$$\ddot{x} + x = 0 \quad (1)$$

Nel piano delle fasi, cioè nel piano spostamento – velocità, il moto della massa è rappresentato da un'orbita ellittica, passante per il punto che rappresenta le condizioni iniziali $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$. Ovviamente, questa orbita corrisponde per il sistema ad una situazione di energia totale costante.

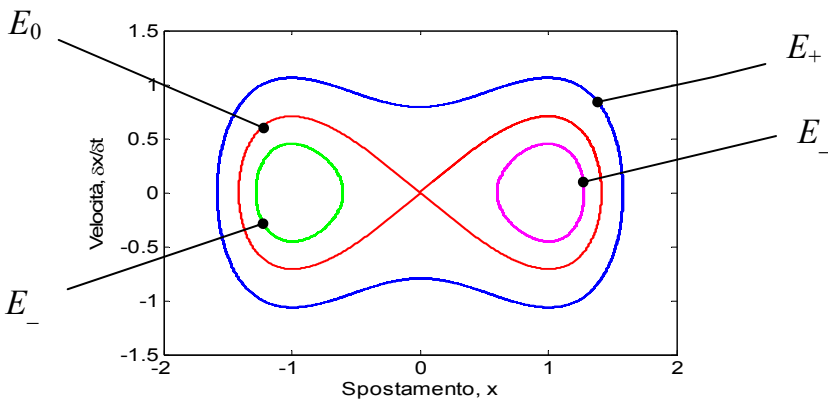


Figura 1 – Traiettorie dell'oscillatore di Duffing nel piano delle fasi.

Nel caso dell'oscillatore di Duffing non smorzato, l'equazione del moto libero è:

$$\ddot{x} - x + x^3 = 0 \quad (2)$$

Sul piano delle fasi, il moto periodico del sistema può dare luogo a quattro tipi di traiettorie (v. fig. 1): il punto $(0, 0)$, la traiettoria E_0 passante per il punto $(0, 0)$, le traiettorie E_- interne ad uno degli "occhi" della traiettoria E_0 e, infine, le traiettorie E_+ esterne alla traiettoria E_0 . Ovviamente, in corrispondenza di una traiettoria l'energia totale del sistema – energia cinetica più energia potenziale elastica – è costante: il punto $(0, 0)$ e la traiettoria E_0 corrispondono a energia nulla, le traiettorie E_+ corrispondono a energia positiva e le traiettorie E_- a energia negativa.

L'equazione di una generica di tali traiettorie è:

$$\dot{x}^2 - x^2 + \frac{1}{2}x^4 = C \quad (3)$$

vale a dire:

$$\dot{x} = \pm \sqrt{x^2 - \frac{1}{2}x^4 + C} \quad (4)$$

La costante C è nulla per la traiettoria E_0 , positiva per le traiettorie E_+ , negativa – nel campo $-0.5 < C < 0$ – per le traiettorie E_- .

Passando a considerare un generico sistema smorzato, eccitato da una forzante esterna, l'equazione del moto ha la forma:

$$\frac{1}{\omega_0^2} \ddot{x} + \frac{\delta}{\omega_0} \dot{x} - x + x^3 = a \cos(\omega_0 t + \varphi) \quad (5)$$

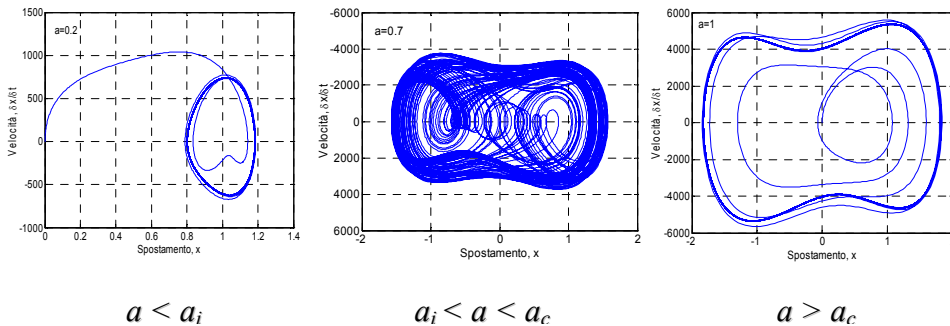


Figura 2 – Comportamento stabile ($a < a_i$), caotico ($a_i < a < a_c$) e stabile ($a > a_c$) dell'oscillatore di Duffing all'aumentare dell'ampiezza a dell'eccitazione.

In questo caso, a parità di valori di δ e di ω_0 , il comportamento dell'oscillatore dipende in modo essenziale dal valore dell'ampiezza a dell'eccitazione. Partendo da condizioni iniziali generiche, per es. nulle ($x(0) = 0, \dot{x}(0) = 0$), dopo il transitorio iniziale, per a minore di un valore critico a_i il sistema tende ad una traiettoria stabile simile alle traiettorie E_- di fig. 1; per a superiore al valore critico a_i il sistema ha un comportamento caotico; e infine, per a superiore ad un secondo valore critico a_c , esso tende ad una traiettoria stabile simile alle traiettorie E_+ di fig. 1 (v. fig. 2).

Facciamo rilevare che i risultati riportati nelle figure 1 e 2 e nelle figure successive sono ottenuti per integrazione numerica: l'osservazione diretta mostra che il metodo e il passo d'integrazione influenzano, sia pure di poco, i risultati e, in particolare, i valori critici a_i e a_c . Per confrontare fra loro i risultati relativi a casi diversi, è pertanto opportuno utilizzare sempre lo stesso metodo e lo stesso passo d'integrazione.

L'aspetto più interessante della transizione da un tipo di comportamento all'altro è il fatto che essa non è graduale, ma pressoché improvvisa: se a passa da un valore poco inferiore ad uno poco superiore ad a_c , l'oscillatore passa "di colpo" dal comportamento caotico al comportamento stabile, rivelando così il pur piccolissimo incremento subito da a .

Il metodo proposto per rivelare la presenza o meno di un segnale di frequenza nota "nascosto" in un segnale sperimentale con alto livello di "rumore" si basa appunto su questo particolare comportamento.

Dalle ultime osservazioni fatte discende immediatamente il procedimento che si può utilizzare per rilevare un segnale di frequenza nota contenuto in un segnale generico.

Si consideri infatti l'equazione:

$$\frac{1}{\omega_0^2} \ddot{x} + \frac{\delta}{\omega_0} \dot{x} - x + x^3 = a \cos(\omega_0 t + \varphi) + s(t) \quad (6)$$

dove $s(t)$ è il segnale da rilevare, ω_0 è la sua pulsazione, a e φ sono l'ampiezza e la fase dell'eccitazione di riferimento; supporremo che sia

$$s(t) = s \cos(\omega_0 t) \quad (7)$$

Se a è poco inferiore al valore critico a_c , il sistema ha un comportamento caotico quando $s = 0$, mentre se $s \neq 0$ ha un comportamento non caotico: in questo caso, infatti, l'eccitazione applicata ha intensità $a + s > a_c$.

3. RILEVAMENTO DEL SEGNALE

Se la situazione è quella rappresentata dalla (6) e il segnale $s(t)$ ha la forma (7), cioè se oltre all'eccitazione di riferimento è presente solo un segnale sinusoidale puro, con la stessa frequenza e la stessa fase del segnale di riferimento stesso, il comportamento dell'equazione (6) è perfettamente identico a quello della (5), ed il risultato è quello rappresentato in figura 2. Poiché, però, un segnale rilevato sperimentalmente non può avere le caratteristiche sopra riportate, è necessario esaminare l'effetto sul comportamento del sistema dei vari fattori presenti nel segnale stesso.

Rumore casuale

Se il segnale da rilevare è accompagnato da un rumore casuale $n(t)$:

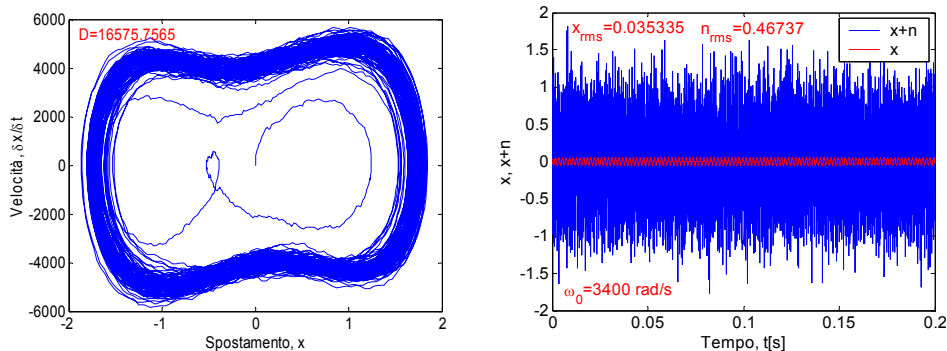


Figura 3 – Segnale sinusoidale in presenza di rumore elevato e risposta dell'oscillatore di Duffing.

$$\frac{1}{\omega_0^2} \ddot{x} + \frac{\delta}{\omega_0} \dot{x} - x + x^3 = a \cos(\omega_0 t + \varphi) + s(t) + n(t) \quad (8)$$

la traiettoria del sistema nel piano delle fasi assume un andamento irregolare, ma rimane della stessa tipologia che avrebbe con rumore nullo: in altre parole, fino a valori anche molto elevati, il rumore non influenza la transizione da comportamento caotico a comportamento regolare.

In figura 3 è riportato il risultato relativo al minimo rapporto segnale/rumore per il quale è stato possibile osservare la transizione caos-non caos, e quindi rilevare la presenza del segnale stesso: si tratta di un segnale sinusoidale immerso in un rumore gaussiano, con rapporto dei valori rms dei due segnali pari a 13.2; le condizioni iniziali sono sempre $x(0) = 0$, $\dot{x}(0) = 0$.

Valutazione del caos

Le due situazioni caos – non caos possono essere giudicate osservando gli andamenti delle traiettorie nel piano delle fasi: gli andamenti sono sempre del tipo di quelli rappresentati in figura 2, e pertanto la differenza è subito evidente.

Per l'esame sistematico di un sistema, però, questo modo di procedere può essere molto scomodo: per individuare, ad esempio, la frequenza sconosciuta di un segnale di interesse, possono essere necessarie anche centinaia di analisi, ed è quindi necessario poter utilizzare un metodo che permetta di automatizzare il giudizio, esprimendo la situazione di caos o di non caos mediante un numero.

In letteratura si trova suggerito a questo scopo il parametro “entropia” (v. ad es. [12]); nel presente lavoro si propongono due nuovi parametri, indicati brevemente come “varianza V” e “distanza D”.

La varianza V è definita nel modo seguente.

In condizioni di caos, il diagramma delle fasi ottenuto dal modello di Duffing presenta traiettorie di ampiezze diverse e con numerosi cappi. In tale situazione, il valore assoluto della derivata della velocità rispetto allo spostamento, $\partial \dot{x} / \partial x$, tende all'infinito per valori dello spostamento x compresi all'interno del suo campo di variabilità e non addensati attorno ai

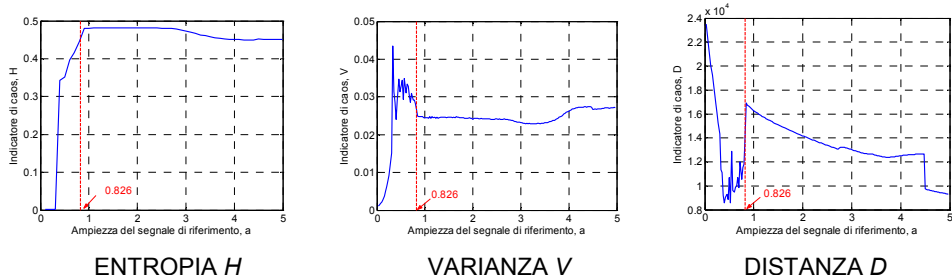


Figura 4 – Indicatori di caos: entropia H , varianza V , distanza D in funzione dell'ampiezza a dell'eccitazione di riferimento. Il valore critico è $a_c = 0.826$.

suoi estremi, come avverrebbe nel caso di assenza di caos. Fissata una soglia inferiore, i_d , per i valori assoluti di tale derivata:

$$\left| \frac{\partial \dot{x}}{\partial x} \right| \geq i_d, \quad (9)$$

si considera la distribuzione dell'evento in funzione dello spostamento e se ne calcola la varianza V . Per meglio evidenziare la presenza del caos, viene eliminata una parte della distribuzione corrispondente alla situazione di non caos, ossia la parte prossima agli estremi del campo di variabilità dello spostamento. Così facendo, si ottiene che il valore della varianza della distribuzione ridotta in presenza di caos sia maggiore di quello corrispondente alla situazione di non caos.

La distanza D è semplicemente la somma dei quadrati delle distanze dall'origine $(0, 0)$ dei punti (x, \dot{x}) delle traiettorie: a parità di altre circostanze, D risulta maggiore nella situazione di moto regolare e minore nella situazione di caos.

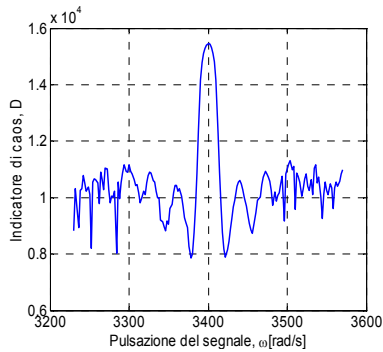
Entrambi i parametri V e D sono dimostrati più efficaci dell'entropia; nel seguito si è sempre adottato il parametro D , che è apparso più comodo ed efficace.

Nella figura 4 sono riportati gli andamenti dei tre parametri – entropia, varianza, distanza – al variare dell'ampiezza a del segnale di riferimento, che in questo caso era l'unico presente: si vede che il valore critico $a_c = 0.826$ è individuato con ottima precisione sia da V sia da D .

Influenza della frequenza e della fase

La transizione caos – non caos è influenzata dalla frequenza e dalla fase del segnale $s(t)$. Questi effetti, già osservati da altri [7 – 13], sono qui evidenziati in termini dell'indicatore di caos D .

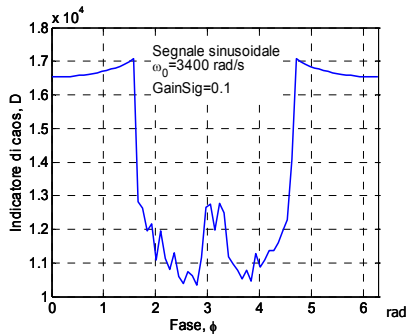
Per quanto riguarda la frequenza, in linea di principio l'oscillatore, il segnale di riferimento $a(t)$ e il segnale da rilevare $s(t)$ dovrebbero avere la stessa pulsazione ω_0 . In realtà, la transizione ha luogo anche se la pulsazione di $s(t)$ è leggermente diversa da quella di $a(t)$. Ricavando l'indicatore D al variare della frequenza di $s(t)$ nell'intorno di quella del segnale di riferimento $a(t)$, si sono ottenuti andamenti del tipo di quello riportato, a titolo di esempio, nella figura 5. Si vede che l'indicatore D assume valori significativamente elevati – cioè



Dati: $T = 0.2$ s, $dt = 0.00001$ s, $\text{GainSig} = 0.5$, $\text{GainNoise} = 0$,
 $\text{AmpSin} = 0.825$, $d = 0.5$, $\text{fase} = 0$, $\omega_0 = 3400$ rad/s.

Figura 5 – Indicatore di caos D in funzione della frequenza del segnale da rilevare. La figura mostra il comportamento dell'oscillatore al variare della pulsazione ω del segnale esaminato, nell'intorno della pulsazione ω_0 del segnale di riferimento, in condizioni di transizione caos \rightarrow non caos. I dati sono: T = durata totale del segnale; dt = passo dell'integrazione numerica; GainSig = ampiezza del segnale da rilevare; GainNoise = ampiezza del rumore; AmpSin = ampiezza del segnale di riferimento, d = fattore di smorzamento.

corrispondenti a caos – anche quando la frequenza del segnale $s(t)$ è nell'intorno di circa più o meno l'1% di quella del segnale di riferimento.



Dati: $T = 0.2$ s, $dt = 0.00001$ s, $\text{GainSig} = 0.1$, $\text{GainNoise} = 0$,
 $\text{AmpSin} = 0.825$, $d = 0$, $\omega_0 = 3400$ rad/s.

Figura 6 – Indicatore di caos D in funzione della fase del segnale da rilevare. La figura mostra il comportamento dell'oscillatore al variare della fase del segnale esaminato, in condizioni di transizione caos \rightarrow non caos. Il significato dei simboli è lo stesso di figura 5.

Per quanto riguarda la fase, in linea di principio il segnale di riferimento $a(t)$ e il segnale da rilevare $s(t)$ dovrebbero avere la stessa fase: solo così, infatti, i due segnali si sommano per dare luogo ad un segnale risultante di ampiezza pari alla somma delle ampiezze dei due segnali; in realtà, però, la transizione ha luogo anche se la fase di $s(t)$ è sensibilmente diversa da quella di $a(t)$. Calcolando l'indicatore D per diversi valori della fase di $s(t)$ rispetto a quella del segnale di riferimento $a(t)$, si ottengono andamenti del tipo di quello riportato, a titolo di esempio, nella figura 6. Si vede che l'indicatore D assume valori significativamente elevati – cioè corrispondenti a caos – anche quando la fase del segnale $s(t)$ è diversa da quella del segnale di riferimento, purché la differenza dia abbastanza lontana da 180° .

4.APPLICAZIONI

Segnale simulato: danneggiamento sull'anello interno di un cuscinetto a sfere

Per verificare la possibilità di impiegare l'oscillatore di Duffing per il monitoraggio e la diagnostica delle macchine, è stato imposto ad un opportuno oscillatore il segnale corrispondente ad un danno sull'anello interno di un cuscinetto a sfere con le caratteristiche riportate in Tabella 1.

Il segnale è stato ottenuto con un modello implementato in LabView®.

Come mostra la figura 7, quando è applicato solo l'eccitazione di riferimento, di ampiezza inferiore all'ampiezza critica a_c , il comportamento è caotico, mentre se si aggiunge anche il segnale che corrisponde al danno sull'anello interno, il comportamento non è più caotico: l'oscillatore è pertanto in grado di rilevare la presenza nel segnale della componente corrispondente al danno da evidenziare.

Lo stesso segnale è riportato in figura 8 nel dominio del tempo (durata totale 1 s) e in quello delle frequenze da 0 a 500 Hz. Sono altresì riportati gli andamenti dello spettro (PSD) nel campo fra 160 e 200 Hz, con risoluzione $\Delta f = 0.978$ Hz e $\Delta f = 5$ Hz. Insieme con questi due andamenti è riportato anche l'andamento dell'indicatore D, calcolato su un segnale della durata di 0.2 s e variando la frequenza del segnale di riferimento da 160 a 200 Hz con passo di circa 0.3 Hz.

Si vede che l'indicatore di caos segnala la presenza della componente cercata con una zona nell'intorno della frequenza caratteristica, nella quale assume valore elevato. L'indicatore stesso risulta meno efficace dell'analisi in frequenza con risoluzione 0.98 Hz, ma un po' più efficace della stessa analisi effettuata con risoluzione di 5 Hz.

Tabella 1 – Caratteristiche del cuscinetto a sfere e del danno

Numero di sfere $N = 12$
Angolo di contatto $\beta = 6^\circ$
Rapporto diametro sfera/Diametro medio $\lambda = 0.180$
$A_i = 1.70$ U.A.(primo caso)
$A_i = 2.50$ U.A. (secondo caso)
Noise = 10 U.A.
Frequenza caratteristica del danno $f_i = 176.9$ Hz

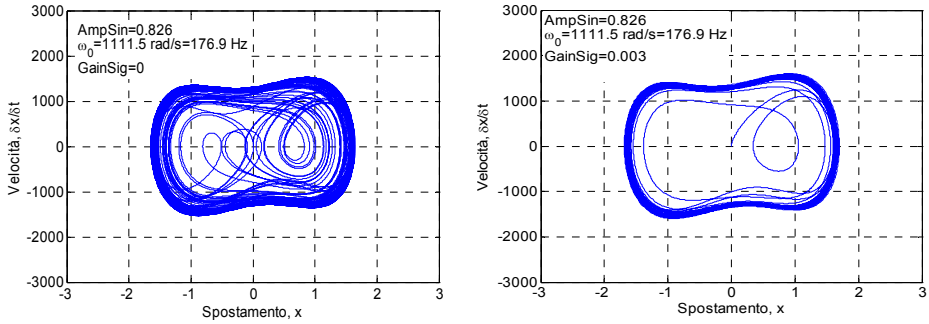


Figura 7 – Traiettorie dell'oscillatore di Duffing nel piano delle fasi, in assenza (a sinistra) e in presenza (a destra) del segnale che simula quello rilevato da un cuscinetto a sfere con un danno superficiale sulla pista di rotolamento dell'anello interno.

Osserviamo per inciso che il metodo impiegato, cioè il calcolo dell'indicatore D in un certo campo di frequenza, permette in sostanza di riconoscere la presenza nel segnale di componenti comprese nel campo esaminato. È da rilevare che il tempo di calcolo, per ogni ciclo, risulta brevissimo: circa 2.6 ms (hp workstation xw4200 – Intel® Pentium® 4 CPU – 3.40 GB – 1.00 GB RAM).

In figura 9 sono riportati gli stessi diagrammi di figura 8, relativi allo stesso cuscinetto ma con danno simulato $A_i = 2.50$ U.A. Anche per questo caso si possono sostanzialmente ripetere le considerazioni esposte per il caso precedente, confermandole.

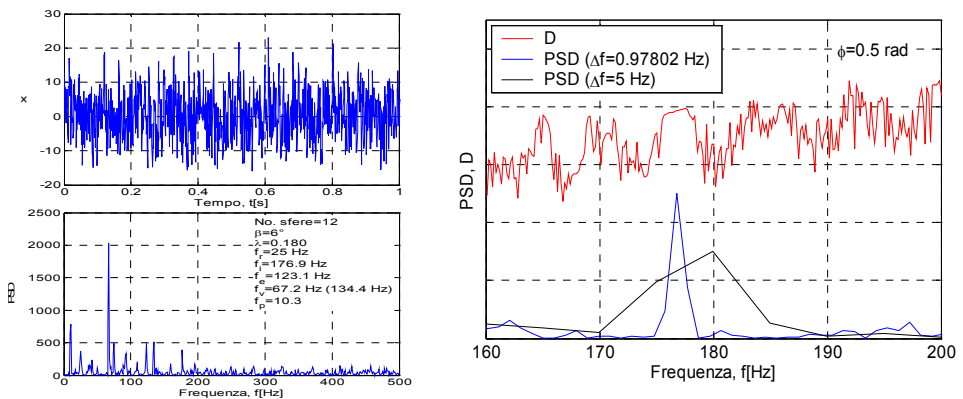


Figura 8 – Segnale che simula quello rilevato da un cuscinetto a sfere con un danno superficiale sulla pista di rotolamento dell'anello interno (primo caso): andamento nel tempo (in alto a sinistra), PSD nel campo 0-500 Hz (in basso a sinistra), PSD con risoluzioni $\Delta f = 0.978$ Hz e $\Delta f = 5$ Hz e andamento dell'indicatore di caos D (a destra).

Tabella 2 – Caratteristiche dell’ingranaggio

Numero di denti del pignone $Z_1 = 28$
Numero di denti della ruota $Z_2 = 55$
Velocità nominale del pignone $n_1 = 1000$ giri/min
Coppia nominale sul pignone: $M_1 = 385$ N·m
<i>Rilevamento in direzione assiale</i>
Frequenza di campionamento: $f_c = 51200$ Hz

Segnale reale: ingranaggio

Per verificare la possibilità di analizzare un segnale reale, cioè rilevato sperimentalmente, si è considerato il segnale relativo ad un ingranaggio con le caratteristiche riportate il Tabella 2.

In figura 10 è riportato l’andamento nel tempo del segnale rilevato e il risultato dell’analisi in frequenza in corrispondenza della frequenza di rotazione ($f_r \approx 16.59$ Hz) e della frequenza di ingranamento ($f_i \approx 464.4$ Hz); nell’intorno della frequenza di ingranamento si notano le bande laterali alle frequenze $f_i \pm f_r$.

Il segnale sperimentale è stato introdotto nell’oscillatore di Duffing, facendone variare la frequenza per intervalli di 0.3 Hz nel campo da 454 a 477 Hz, e calcolando ogni volta l’indicatore di caos D. I risultati sono riportati in figura 11. Si vede che l’indicatore D individua la presenza del segnale alla pulsazione di interesse $\omega_0 = 2918$ rad/s con una zona in cui esso assume valore elevato, indicando la transizione caos non caos; si nota tuttavia che il risultato dipende anche dalla fase del segnale di base: in questo caso, l’individuazione è più precisa assumendo la fase pari a 1.5 rad rispetto al caso di fase nulla.

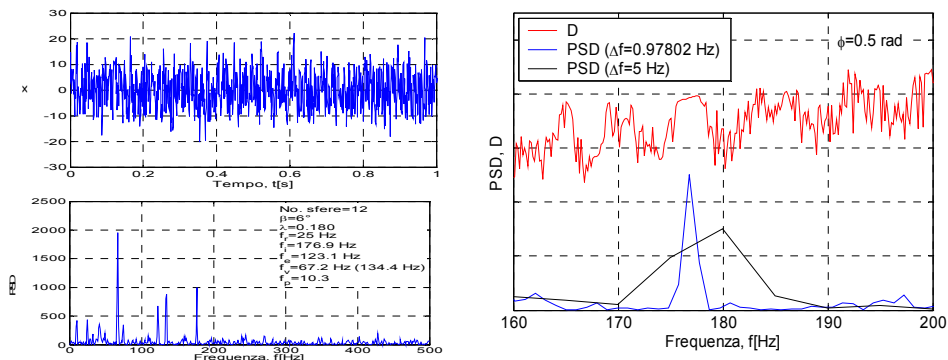


Figura 9 – Segnale che simula quello rilevato da un cuscinetto a sfere con un danno superficiale sulla pista di rotolamento dell’anello interno (secondo caso): andamento nel tempo (in alto a sinistra), PSD nel campo 0-500 Hz (in basso a sinistra), PSD con risoluzioni $\Delta f = 0.978$ Hz e $\Delta f = 5$ Hz e andamento dell’indicatore di caos D (a destra).

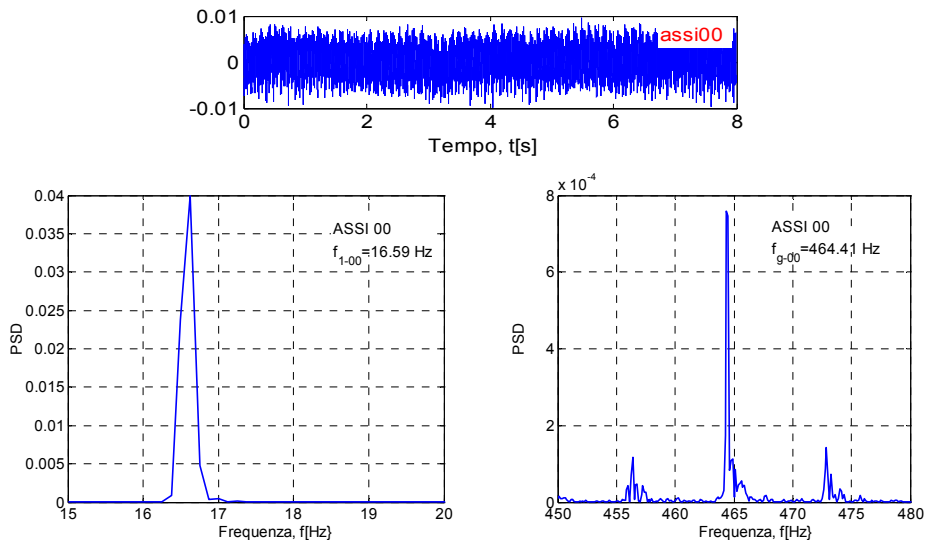


Figura 10 – Segnale rilevato sperimentalmente su un ingranaggio con le caratteristiche riportate nella Tabella 2 (in alto) e PSD del segnale stesso nel campo 0-20 Hz (in basso a sinistra) e nel campo 450-480 Hz (in basso a destra).

Come nel caso del segnale simulato di figura 9, si vede che rispetto all'analisi in frequenza il risultato dell'indicatore di caos è meno accurato se l'analisi viene effettuata su un segnale relativamente lungo ($\Delta f = 0.125$ Hz, durata del segnale 8 s), mentre è più accurato se l'analisi viene effettuata su un segnale di breve durata ($\Delta f = 5$ Hz, durata del segnale 0.2 s).

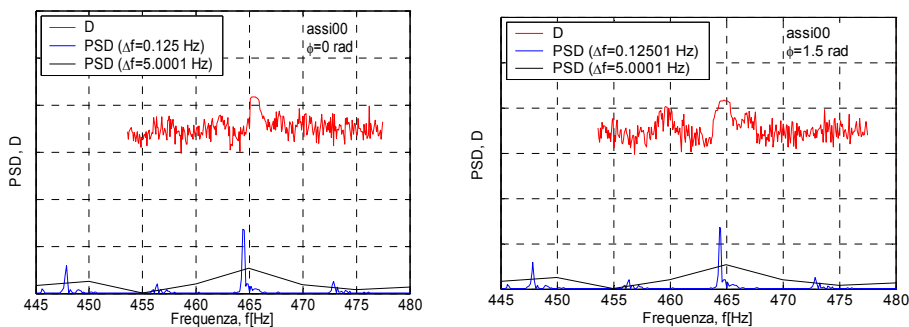


Figura 11 – Analisi del segnale sperimentalmente riportato in fig. 10: confronto fra indicatore di caos D, PSD con risoluzione $\Delta f = 0.125$ Hz e PSD con risoluzione $\Delta f = 5$ Hz e andamento dell'indicatore di caos D (a destra), nell'intorno della frequenza di ingranamento. Il calcolo dell'indicatore D è stato effettuato per due valori della fase ϕ .

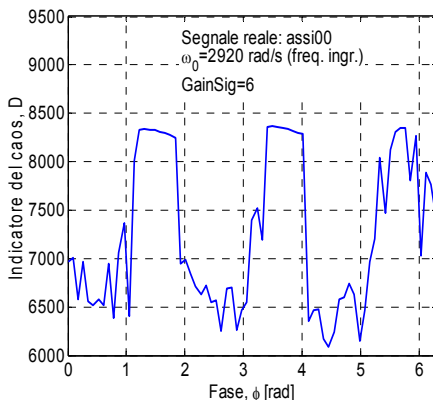


Figura 12 – Andamento dell’indicatore di caos D in funzione della fase del segnale di riferimento nell’analisi del segnale sperimentale dell’ingranaggio di Tabella 2; v. figure 10 e 11.

In figura 12 è riportato l’andamento dell’indicatore di caos D alla frequenza di ingranamento, in funzione della fase dell’eccitazione di riferimento. Viene confermata l’influenza di tale fase, che è collegata alla natura del procedimento e ne costituisce un evidente aspetto negativo.

5. CONCLUSIONI

È stata verificata l’effettiva possibilità di impiegare l’oscillatore di Duffing per rilevare la presenza di una debole componente sinusoidale in un segnale anche molto rumoroso. Il metodo può trovare applicazione, per es., nel monitoraggio e nella diagnostica industriali, per la sorveglianza dello stato di salute o per il controllo di qualità delle macchine o di loro componenti.

È stato altresì proposto un nuovo indicatore di caos, che è risultato efficace. L’impiego dell’indicatore è praticamente indispensabile per automatizzare indagini sistematiche. A titolo di esempio, si sono esaminati due segnali simulati di cuscinetti a rotolamento con un danno sulla pista di rotolamento dell’anello interno e il segnale reale rilevato su un ingranaggio. Il metodo è risultato efficace per il rilevamento della componente cercata, che in tutti i casi esaminati era debole ed immersa in un segnale molto rumoroso.

Peraltro, non sono emerse particolarità che permettano di giudicare l’impiego dell’oscillatore di Duffing più efficace degli altri metodi comunemente in uso, a cominciare dalla semplice analisi in frequenza. Solo nel caso di tempi di acquisizione molto piccoli la presenza di una componente di piccola intensità è stata rilevata meglio con l’oscillatore di Duffing che con l’analisi di Fourier.

Il procedimento ha comunque dato risultati corretti, per cui può essere opportuno approfondirne lo studio per ricercarne altre eventuali caratteristiche positive, finora peraltro non evidenziate.

BIBLIOGRAFIA

- [1] Collacott R. A., 1979, *Vibration Monitoring and Diagnosis*, J. Wiley & S., ISBN 0 7114 5201 6
- [2] Lipovszky G., Sòlimvari K., Varga G., 1990, *Vibration Testing of Machines and their Maintenance*, Elsevier, ISBN 0 444 98808 4
- [3] Williams J. H., Davies A., Drake P. R., 1994, *Condition-based Maintenance and Machine Diagnostics*, Chapman & Hall, ISBN 0 412 46500 0
- [4] Dalpiaz, G. et al., 2000, "Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears", *Mechanical Systems and Signal Processing* **14**, pp. 387-412
- [5] Rubini, R., Meneghetti, U., 2001, "Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings", *Mechanical Systems and Signal Processing*, **15**, pp. 287-302
- [6] Mobley R. K., 2002, *An Introduction to Predictive Maintenance*, Butterworth Heinemann, ISBN 0 7506 7531 4
- [7] Wang, G. et al., 1999, "The application of chaotic oscillators to weak signal detection", *IEEE Transactions on Industrial Electronics*, **46** (2), pp. 440-444
- [8] Wang, G. et al., 2002, "Estimation of amplitude and phase of a weak signal by using the property of sensitive dependence on initial conditions of a nonlinear oscillator", *Signal Processing*, **82**, pp. 103-115
- [9] Wang, G., He, S., 2003, "A quantitative study on detection and estimation of weak signals by using chaotic oscillators", *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, **50**, pp. 945-953
- [10] Le B. et al., 2005, "Chaotic oscillator and other techniques for detection of weak signals", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **E88-A** (10), pp. 2699-2701
- [11] Li C., Qu L., 2005, "A nonlinear diagnosis method of gear early fatigue crack", in IEEE International Conference on Vehicular Electronics and Safety, pp. 134-139 ISBN 0-7803-9435-6
- [12] Li C., Qu L., 2007, "Applications of chaotic oscillator in machinery fault diagnosis", *Mechanical Systems and Signal Processing*, **21**, pp. 257-269
- [13] Liu D. et al., 2005, "Weak signal detection based on chaotic oscillator", in Fourtieth IAS Annual Meeting Conference, pp. 2054- 2058, ISBN 0-7803-9208-6
- [14] Brown, R. et al., 1996, "Is sensitive dependence on initial conditions nature's sensory device?", *Int. J. Bifurc. Chaos* , **2**(1), pp. 193-199
- [15] Jordan, D. W., P. Smith, 2007, *Nonlinear Ordinary Differential Equations*, 4th ed., Oxford University Press, ISBN 978 0 19 920824 1

INTRODUCTION TO A SIMPLY AND FAST ALGORITHM FOR VARIABLE SPEED BEARING DIAGNOSTICS

Marco Cocconcelli

*Department of Engineering Sciences and Methods,
University of Modena and Reggio Emilia, Italy
E-mail: marco.cocconcelli@unimore.it*

Riccardo Rubini

*Department of Engineering Sciences and Methods,
University of Modena and Reggio Emilia, Italy
E-mail: riccardo.rubini@unimore.it*

Abstract. *In this paper a new algorithm for variable speed bearing diagnostics is proposed. The diagnostics of ball-bearing is used to be focused on constant speed applications. In fact at constant speed bearing faults cause a vibration with particular characteristics that the analysis in the frequency domain reveals. The geometry of the bearing parts allows to find the correlation between faults frequencies and the rotational frequency of the shaft on which the bearing is mounted on. If the shaft moves with variable speed or motion inversion, the instantaneous rotation frequency changes continuously and no fault characteristic frequency are available. Under the hypothesis that the shaft position profile is known, this paper suggests to perform the diagnostics on variable speed applications in two steps. In the first one both bearing geometry and the shaft position profile are used to make a simulated fault signal which is expected if a specific fault would be present, while in the second step a comparison between the simulated signal and the real vibration signal is made in order to verify the presence of the fault in the bearing. Different test-studies are reported in order to validate the algorithm.*

Keywords: *variable speed, diagnostics, correlation, bearing faults*

1. INTRODUCTION

Predictive maintenance is a challenge that it is increasing its worth in manufacturing industries. In fact it joints two common needs: on one side it allows to plan the replace of specific machinery's components to avoid sudden and unexpected downtime, on the other side it can suggest the parts replacement only if the component is really damaged, decreasing the maintenance costs. In particular the most frequent cause of failure in the electric motors is the damage of the ball-bearing. The classic literature methods are mainly focused on constant speed applications [1–4]. In fact it is well known that if a damage occurs to the bearing it could be revealed by the frequency analysis of the motor vibration signal. The presence of a fault, e.g. on the outer race, produces an impact every time a sphere passes on it and if the

shaft is moving with a given rotational frequency f_r than the fault makes a train of impulses with a characteristic frequency which is proportional to f_r . This proportional coefficient is calculated once the bearing geometry is known. In particular the ball-bearing can be modelled as a epicyclic gear where the races and the spheres act like gears and the cage like the carrier [2, 3], then the relative velocities between all bearing components can be calculated and they only depend on bearing geometry. Unfortunately the classic techniques cannot be applied to perform diagnostics on modern direct-drive motors. These motor are real-time controlled in order to follow complex shaft position profiles with variable speeds, avoiding the presence of a gear reducer between the motor and the user. In this case there is no more a constant rotational frequency f_r and then there are no more fault characteristic frequencies in the signal spectrum. The first attempt to analyse variable speed applications was the Order Tracking (OT) [1, 3]. The OT passes from a time sample domain to an angular sample domain, that is it takes the shaft angular displacement as reference system and it collects a given number of sample in shaft revolution regardless of the shaft speed. The acquired signal is then analysed in “order domain” which stands at spatial sampling as the frequency domain to time sampling. The main disadvantages of the OT are the need of a minimum number of revolutions to increase the resolution of the order analysis and the high cost of the hardware needed. The higher costs have been overcome by the Computed Order Tracking (COT) [5, 6] which substitute the hardware with a software manipulation of the data taken in time domain once a tachometer signal is available. Unfortunately the need of a given complete revolutions of the shaft remains and it cannot be always satisfied by the direct-drive motors where there is frequently a motion inversion of the shaft after a single revolution or less. Other algorithms based on time-frequency techniques (e.g. Wavelet Transform) generally require long computational-time and their results are quite complicated and they need qualified staff to be read. In this paper we propose a simply and fast algorithm for variable speed diagnostics. We start from two considerations: the kinematic model of ball-bearing holds true regardless of the instant speed of the shaft; the control loop of the electric motor require an internal encoder to determine the position of the shaft at each time instants and this position signal could be carried outside the machine. The position profile of the motor is used together with the bearing kinematic model to make an artificial vibration signal which simulates the presence of a bearing fault in a motor which is rotating with the given position profile. Indeed different type of faults (e.g. inner or outer race) produces different simulated signals. Each simulated signal is then compared with the real acquired vibration signal. A cross-correlation between real and simulated signal is done and the output is analysed in order to verify the presence of the presumed fault. The flowchart of the process is shown in Fig.1 The aim of this procedure is to quickly analyse data on a quality level, which could be act as an alarm signal for the user who shall made a deep-in time-consuming analysis only in few cases. The paper is structured as follow: in Section two the algorithm for the format of the artificial faulted signal is explained. The comparison between real and simulated signal is the subject of the Section three. At least, Section four shows the application of the procedure to some test-cases and real cases.

2. ARTIFICIAL FAULT FORMAT

In this Section an artificial signal which simulates a fault on bearing is made. This artificial fault algorithm accepts as inputs an arbitrary motion profile and the geometrical dimensions of the bearing, while it gives the simulated vibration signal for different faults as output.

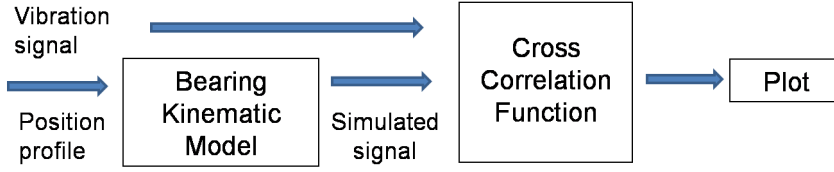


Figure 1: Flowchart of the algorithm

Artificial Fault Algorithm

The kinematics behaviour of the ball bearing is well described by the classic epicyclic gear approximation: the inner race is the sun gear, the bearing' spheres are the planet gears, the outer ring is the frame and the cage of the bearing is the carrier. The sun gear angular velocity is given, because the inner ring of the bearing is mounted on the motor shaft. As a consequence the angular velocity of the whole elements could be calculated and the rate between it and the shaft rotational velocity depends on the geometric characteristics of the bearing. In Eq. (1) a kinematics relation between the angular velocity of bearing elements is reported. Indeed those relations had been found under the hypothesis of no-slip between the bearing' elements.

$$\frac{\omega_n - \omega_p}{\omega_1 - \omega_p} = \tau_o \quad (1)$$

where ω_n , ω_p and ω_1 are the angular velocities of the outer race, the cage and the inner race, respectively and τ_o is the transmission rate between outer and inner races when the cage velocity is zero. This ratio could be expressed as a function of the geometrical dimensions of the bearing. Since $\omega = 2\pi f$, those ratios describe also the relations between the rotational frequency of the shaft and the rotational frequencies of the bearing elements. As an example if a fault is present on the outer race every time a single sphere passes on the fault, it generates an impact with a periodicity which depends on the rotational frequencies of the shaft and the cage and the angular position of the damaged component. These characteristic fault frequencies could be calculated through the Eq. (1) – also by substituting one of the ring velocities with rolling element' spin ω_v – and they are reported in Eqs. (2), (3), (4) and (5).

$$f_e = \frac{Z}{2} f_r (1 - \lambda \cos(\beta)) \quad (2)$$

$$f_i = \frac{Z}{2} f_r (1 + \lambda \cos(\beta)) \quad (3)$$

$$f_c = \frac{1}{2} f_r (1 - \lambda \cos(\beta)) \quad (4)$$

$$f_v = f_r \frac{(1 - \lambda \cos(\beta))^2}{2\lambda} \quad (5)$$

where f_e , f_i , f_c , f_v and f_r are the outer race fault frequency, the inner race fault frequency, the cage fault frequency, the rolling element fault frequency and the rotational frequency,

respectively. Coefficient $\lambda = \frac{d}{D}$ is the ratio between the ball diameter d and the cage mean diameter D , β is the contact angle between balls and bearing races and Z is the number of balls in the bearing. If the angular velocity ω_r is not constant, the fault frequencies are still related to the rotational frequency f_r but they satisfy equations (2) – (5) for each time instant. In this case the better is to move from frequency-domain to time-domain analysis where equations like (2)–(5) are available but they express a relation between angular velocities instead of frequencies and they are reported in Eqs. (6), (7),(8) and (9).

$$\omega_e = 0 \quad (6)$$

$$\omega_i = \omega_r \quad (7)$$

$$\omega_c = \frac{\omega_r}{2}(1 - \lambda \cos(\beta)) \quad (8)$$

$$\omega_v = \omega_r \left(\frac{1 + \lambda - \lambda^2(\cos^2(\beta) + \cos(\beta))}{2\lambda} \right) \quad (9)$$

Equations (6) – (9) between the rotational velocity ω_r and angular rotation of each part of the bearing allow to predict the time instant when an impact occurs – at the contact between bearing components in the presence of a surface failure – for any kind of motion profile $\omega_r(t)$. In fact the knowledge of the velocity profile $\omega_r(t)$ allows to determine the expected vibration signal of a damaged bearing.

Simulation Of The Fault On The Outer Race. Consider a fault on the outer race as example. The outer ring is fixed and each time a ball passes on the fault an impulse is generated. In this case we follow the next steps:

1. Determine the corresponding angular velocity of the cage: Eq. (8)
2. Determine the angular displacement of the cage: $\phi(t) = \int \omega_c(t)dt$
3. Determine the rotational range of the shaft: $\Phi = \phi_{max} - \phi_{min}$
4. Determine the angular pitch between two adjacent balls due to the presence of the cage: $\delta_\phi = \frac{2\pi}{Z}$
5. Determine the number of impacts in the whole displacement range of the cage: $n = \lfloor \frac{\Phi}{\delta_\phi} \rfloor$ where n is the bigger integer minor or equal than $\frac{\Phi}{\delta_\phi}$.
6. Determine all the time instants when the cage rotates of δ_ϕ

Simulation Of The Fault On The Inner Race. Consider now a fault on the inner race as example. In that case the relative velocity between inner ring and cage has to be calculated in order to determine the correct time interval between impacts. The steps to be followed are:

1. Determine the angular velocity of the cage: Eq. (8)
2. Determine the angular velocity of the inner ring: Eq. (7)

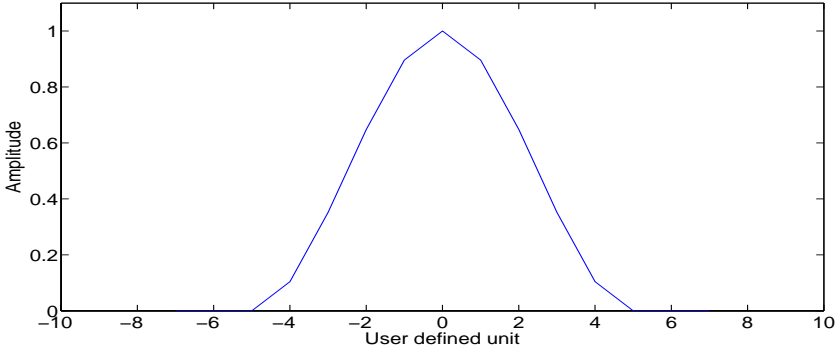


Figure 2: Example of simulated peak

3. Determine the relative angular velocity between the cage and the inner ring: $\omega_{ci}(t) = \omega_c(t) - \omega_i(t)$
4. Determine the angular displacement of the cage: $\phi(t) = \int \omega_{ci}(t)dt$
5. Determine the rotational range of the shaft: $\Phi = \phi_{max} - \phi_{min}$
6. Determine the angular pitch between two adjacent balls due to the presence of the cage:
 $\delta_\phi = \frac{2\pi}{Z}$
7. Determine the number of impacts in the whole displacement range of the cage: $n = \lfloor \frac{\Phi}{\delta_\phi} \rfloor$ where n is the bigger integer minor or equal than $\frac{\Phi}{\delta_\phi}$.
8. Determine all the time instants when the cage rotates of δ_ϕ

Since only the time instants when an impact occurs are relevant, the amplitude is not modulated with the rotational frequency as in real cases.

Similar considerations allow to make the simplified fault model for the cage damage or bearing's ball damage.

In all the different fault cases a n-dimensional zero vector is taken and a peak is added to this zero vector in correspondence of the time instant when an impact occurs according to the specific fault model chosen. The resulting vector has the dimension n equal to the length of the position vector. More simulated signals could be produced, e.g. to simulate the vibration signal of a bearing with two different faults at the same time, etc... The peak added is not a single element impulse, but it has a bell-shape with a fixed even if arbitrary length. An example of single peak is shown in Fig.2

3. COMPARISON BETWEEN ARTIFICIAL AND REAL FAULT SIGNAL

In this Section we compare the artificial fault signals – made in the previous Section – with the signal we want to analyse in order to determine the presence of a specific defect on the bearing. The main idea is the use of an artificial signal as a revealing signal.

Since a given simulated signal is the theoretical expected signal in case of a fault condition,

a function is need which estimates resemblance between simulated and on-study signal: we choose the cross-correlation function.

The mathematical expression of the cross-correlation between functions f and g_j is:

$$C_j = f(t) \star g_j(t) = \int_{-\infty}^{\infty} f(\tau)g_j(t + \tau)d\tau \quad (10)$$

where f is the real vibration signal acquired in the field, g_j is the j -th type of the artificial signal (e.g. outer race fault, inner race fault, etc ...).

The artificial signal could be expressed as a sum of impulses (see Fig.2), each of them happens at different time instants. These time instants, let's call them t_k , are determined through the kinematic model of the bearing as described in the previous Section. The formula of the generic artificial signal g_j becomes:

$$g_j(t) = \sum_k g_{imp}(t - t_k) \quad (11)$$

where g_{imp} is the single impulse function, and the number of time instants t_k depends on the j -th type of fault. Introducing Eq.(11) into Eq.(10), the cross-correlation formula proposed in this paper is given:

$$C_j = \sum_k [f(t) \star g_{imp}(t)]_{(t-t_k)} \quad (12)$$

Pay attention to the nature of the signal we are considering: in the cross-correlation function, the extreme limits of the integral in Eq.(10) moves from $-\infty$ to $+\infty$ while both the signals have a limited length of time. During the calculation mathematical software usually suppose that the signals repeat themselves to infinity. Suppose that Δt is the time interval between two subsequent impacts, unfortunately a generic vibration data record does not contain a integer number of Δt and then errors occur when the data set is repeated in the cross-correlation calculation (call them "cut-and-paste" errors). While the real signal is always affected of cut-and-paste errors, Eq.(12) allows to correct those errors in the simulated signal. In fact, suppose that the first impact occurs at instant t_0 , then the Eq.(12) changes and it is shown in Eq.(13).

$$C_j = \sum_k [f(t) \star g_{imp}(t)]_{(t-t_k-t_0)} \quad (13)$$

If T denotes the data time length and n is the number of impacts in that time, then t_0 is limited between $[0, (T - n\Delta t)]$. We divide this small interval in a given number of parts (e.g. twenty), we perform the evaluation of Eq.(13) for each estimated value of t_0 and then we take the average of all the C_j calculated. This process takes a very few times because the cross-correlation calculus remains the same at every estimation, but only addition operations change. The average function allows to diminish the contribute of C_j when the t_0 value is far from the real value, and moreover it allows to increase the contribute of C_j when the t_0 value is close to the real value.

If the plot of the cross-correlation function shows evident peaks with a slight spread of peaks close to the main one, it means that the bearing is affected by the specific fault to whom the real signal was correlated (e.g. outer race fault). At the same time, the cross-correlation between the real signal and another type of artificial fault (e.g. inner race fault) is expected to

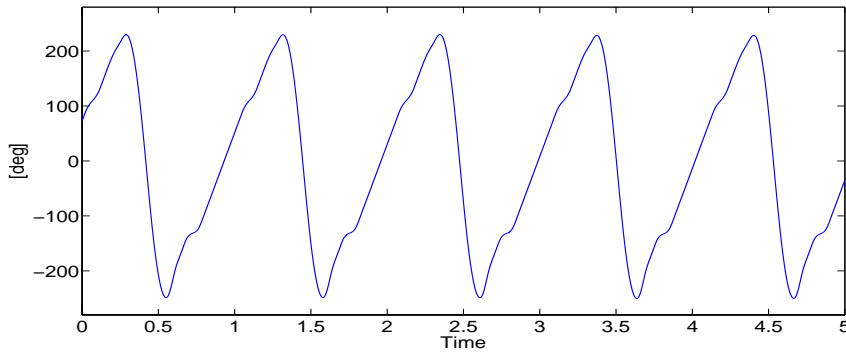


Figure 3: Position profile

Table 1: Main geometrical characteristics of SKF 6309

<i>SKF</i>	6309
d	17.46 [mm]
D	74.54 [mm]
$\lambda = d/D$	0.234
β	8 [deg]
Z	8

be characterised by low amplitude peaks, and significant increase of the peaks' spread close to the main one.

In the next Section, the suggested procedure is applied to an ad-hoc simulated signal and to a naturally damaged bearing.

4. RESULTS

In this Section the proposed procedure is applied to two cases: an ad-hoc simulated signal and a real signal. In both cases the position profile is whose shown in Fig.3 and the bearing is supposed to be damaged on the outer race. In the real case the fault is made with a milling cutting tool.

The bearing used in the real case is a SKF 6309 and its main geometrical characteristics are collected in Tab.1. In the Section 2. the meaning of the letters used had been explained. The sampling frequency is assumed to be 10 kHz and the acquisition time is 5 seconds.

Simulated Signal

The first step to validate the procedure is to create ad-hoc signals with the same algorithm which creates the artificial faulted signals, and test them to look at the procedure response in ideal conditions (i.e. when the on-test signal is the same as the testing one).

The position profile and the bearing characteristic dimensions – see Fig.3 Tab.1 – are used as input of the algorithm. The outputs are the outer-race-fault simulated signal (g_o) and

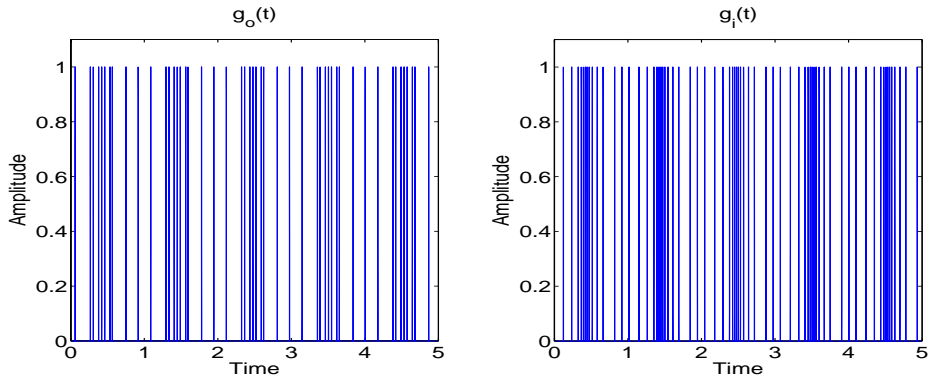


Figure 4: Simulated vibration signal: fault on the outer race (g_o) and on the inner race (g_i)

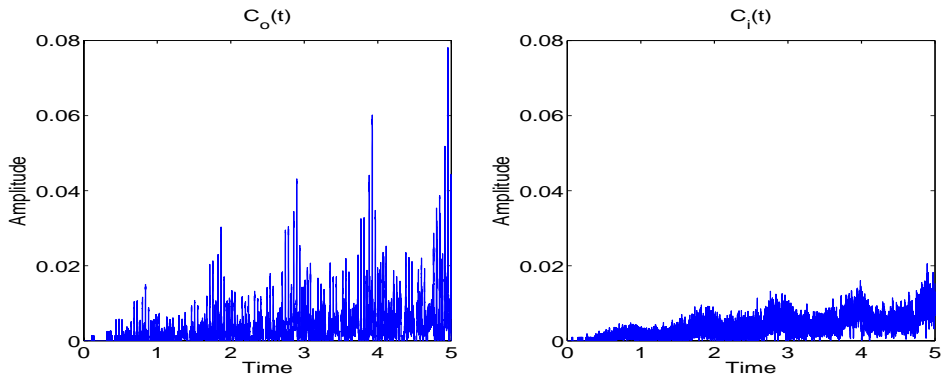


Figure 5: Cross-correlation function: fault on the outer race (C_o) and on the inner race (C_i)

the inner-race-fault simulated signal (g_i) which are shown in Fig.4.

The on-test signal $f(t)$ is assumed to be the outer-race-fault simulated signal: $f(t) = g_o(t)$. Then the cross-correlation functions between $f(t)$ and fault simulated signal are done by Eq.(13) and they are named $C_o(t)$ and $C_i(t)$ in the case of outer and inner race fault respectively. The cross-correlation functions $C_o(t)$ and $C_i(t)$ are shown in Fig.5.

As expected, $C_o(t)$ shows high peaks with steep slope of the secondary peaks amplitude, while $C_i(t)$ realizes the periodicity of the signals is the same, but it does not find any correlation between the expected fault on the inner race and the on-test signal. The presence of some peaks with relevant amplitude in the $C_o(t)$ plot may be probably due to the “cut-and-paste” errors in the cross-correlation calculation.

Real Signal

In this second step the procedure is applied to a vibration signal acquired from a real outer-ring-faulted bearing. In particular the SKF 6309 ball bearing is mounted on a electric servo-

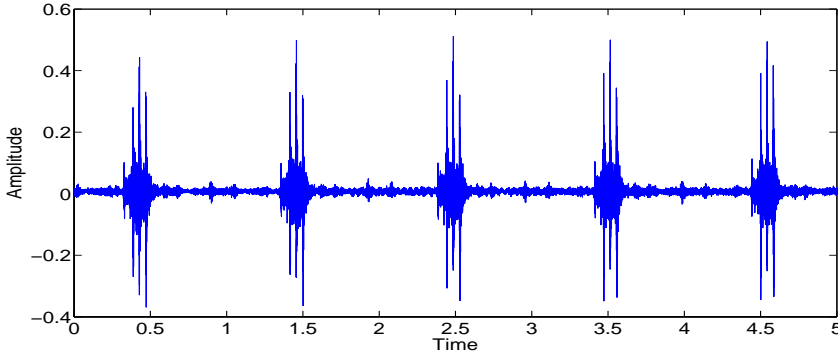


Figure 6: Vibration signal of bearing faulted on the outer race

motor brushless which is drive-controlled to perform the position profile shown in Fig.3. A piezoelectric accelerometer is mounted on the motor surface close to the bearing seat. The acquired vibration signal $f(t)$ is shown in Fig.6.

In that case the on-test signal $f(t)$ is compared with the artificial fault signals $g_o(t)$ and $g_i(t)$ which are the same used in the previous example and shown in Fig.4. The results of the cross-correlation function between those signals are reported in Fig.7.

In the real case the two results are more similar but important differences hold on. From a quantitative point of view the peaks of $C_o(t)$ are higher than ones of $C_i(t)$. From a qualitative point of view the peaks of $C_o(t)$ are more defined, that is there is a main peak and the narrow band shows small amplitudes if compared to the main one. In $C_i(t)$ instead there is a spread of peaks amplitude that is close to the main one. Those characteristics suggest that there is a low correlation between $f(t)$ and the inner-race-fault simulated signal $g_i(t)$, while there is a good and acceptable correlation between $f(t)$ and the outer-race-fault simulated signal $g_o(t)$.

Pay attention to the β value, it is the angle between the real and ideal contact line in the bearing. More precisely the contact line is specified by the contact points between the spheres and the outer and inner races. Due to clearance in the bearing and to the mounting operations, the β value changes from an ideal value of 0 deg to a real value which in literature is set in the range $[0, 10]$ deg. Unfortunately this parameter cannot be exactly determined. In Tab.1 the β value is just estimated for the real case while it can be set in the simulated case. To increase the performances of the proposed procedure, the user could run the algorithm another time, changing the β value.

5. CONCLUSIONS

In this paper a new method to perform diagnostics on variable speed applications is proposed. It requires as input the position profile of the motor, the vibration signal and the geometrical characteristics of the bearing, and it gives as output the cross-correlation function between the vibration signal and specific artificial signals. These signals simulate the expected vibration signal when a specific fault is present. This technique has been proved on two test-cases. In the first one an ad-hoc simulated signal is used to investigate the algorithm response in an ideal case, while in the second one a bearing is naturally damaged and its vibration signal is

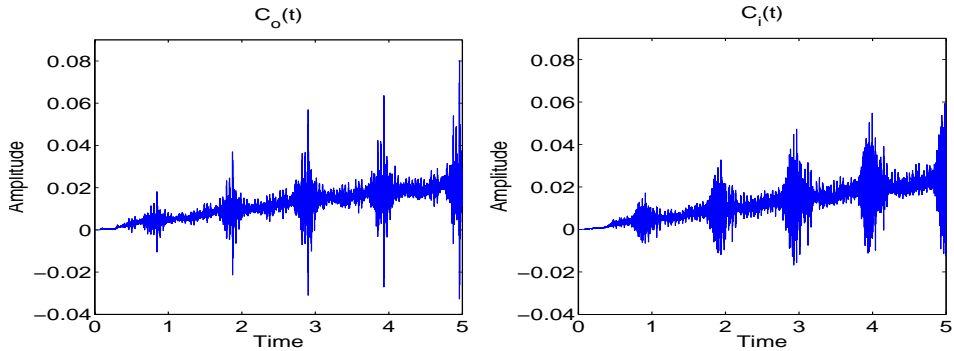


Figure 7: Cross-correlation function: fault on the outer race (C_o) and on the inner race (C_i)

used as on-field test. In both cases the results are good and promising, the procedure allows to correctly predict the specific fault present. The proposed procedure should be improved with a pre-processing of the vibration data in order to denoise it, trying to extract specific failure components of the signal. This step will be analysed by the authors in a future work.

REFERENCES

- [1] Taylor, J. I., 2003. *The Vibration Analysis Handbook*. VCI.
- [2] Taylor, J. I., and Kirkland, D. W., 2004. *The Bearing Analysis Handbook: A Practical Guide for Solving Vibration Problems in Bearings*. VCI.
- [3] Harris, C. M., and Piersol, A. G., 2001. *Harris' shock and vibration handbook*. McGraw-Hill Professional.
- [4] Berry, J., 1991. "How to track rolling element bearing health with vibration signature analysis". *Sound and Vibration*, **25**, pp. 24–23.
- [5] Potter, R., 1990. "A new order tracking method for rotating machinery". *Sound and Vibration*, **24**, pp. 30–34.
- [6] Potter, R., 1990. Tracking and resampling method and apparatus for monitoring the performance of rotating machines.

COLLABORAZIONI FRA L'UNIVERSITÀ DEGLI STUDI DI PARMA E CENTRI DI RICERCA INTERNAZIONALI NELL'AMBITO DELLA MECCANICA APPLICATA

Edzeario Prati
*Dipartimento di Ingegneria Industriale ,
Università degli Studi di Parma, Italy
E-mail: prati@ied.unipr.it*

Alessandro Tasora
*Dipartimento di Ingegneria Industriale ,
Università degli Studi di Parma, Italy
E-mail: tasora@ied.unipr.it*

Marco Silvestri
*Dipartimento di Ingegneria Industriale ,
Università degli Studi di Parma, Italy
E-mail: silvestri@ied.unipr.it*

Abstract. *La presente memoria intende presentare, in forma sintetica, attività di ricerca originate da collaborazioni internazionali fra il gruppo di Meccanica Applicata alle Macchine dell'Università degli Studi di Parma e centri di ricerca esteri. Tali indagini scientifiche si collocano essenzialmente entro le quattro aree tematiche dell'automazione, della meccanica computazionale, della tribologia e della meccanica delle vibrazioni.*

Keywords: *tribologia, multibody, robotica, automazione, meccanica computazionale.*

1. INTRODUZIONE

La scienza moderna, sia nella forma più teorica ed astratta, sia in una declinazione tecnologica ed applicativa come nel caso dell'ingegneria meccanica, trae evidente giovamento dallo scambio di risorse, conoscenze ed opinioni fra studiosi di diverse nazionalità.

Orientato da tale considerazione, nel corso degli ultimi anni il gruppo di Meccanica Applicata alle Macchine dell'Università di Parma ha inteso rafforzare i propri rapporti con centri di ricerca esteri. L'intensificato rapporto con ricercatori stranieri è stato veicolo di risultati innovativi e vantaggiosi per tutti i soggetti coinvolti, dimostrando l'utilità della coesione di molteplici studiosi attorno a progetti di ampio respiro e con orizzonti internazionali.

Nel presente articolo si descriveranno i risultati più significativi di tali attività, con particolare riferimento a quattro aree tematiche che trovano naturale collocazione nell'alveo

della Meccanica Applicata alle Macchine, ovvero la meccanica computazionale, l'automazione, la tribologia e la meccanica delle vibrazioni.

2. COLLABORAZIONE CON ARGONNE NATIONAL LABS

Il centro di ricerca Argonne National Labs (ANL), sito in Illinois e gestito dalla vicina Università di Chicago, rappresenta il primo laboratorio degli Stati Uniti ed è internazionalmente riconosciuto come uno dei più importanti centri di ricerca al mondo. È stato fondato nel secolo scorso da Enrico Fermi e da sempre è autorevole sede di ricerche nell'ambito del supercalcolo, della matematica applicata, della fisica e dell'ingegneria, con particolare attenzione alle applicazioni legate all'energia nucleare.

Preesistenti attività di ricerca dell'Ing. Tasora presso l'Università degli studi di Parma, in merito a metodi di simulazione multibody, nel 2007 hanno motivato l'instaurarsi del rapporto di collaborazione con il dipartimento di scienze matematiche e computazionali (MCS) dei laboratori Argonne, in particolare col Prof. M. Anitescu.

La ricerca mira a due obiettivi. Il primo, di carattere teorico, riguarda lo sviluppo di metodi numerici innovativi per la simulazione di sistemi multi corpo con elevato numero di contatti con attrito; a tal fine si è sviluppato un algoritmo fondato sulla soluzione iterativa di problemi complementari conici convessi (CCP) che ha dimostrato notevoli doti di efficienza e robustezza.

Il secondo obiettivo concerne l'applicazione del metodo numerico sopracitato al fine di simulare il flusso granulare del combustibile all'interno di un reattore nucleare di IV generazione (v. Fig.1). Tale reattore, definito Pebble Bed Reactor (PBR), afferisce alla categoria VHTR e rappresenta una strada promettente per la realizzazione di impianti di quarta generazione, data la sua sicurezza intrinseca, l'alta efficienza e la possibilità di

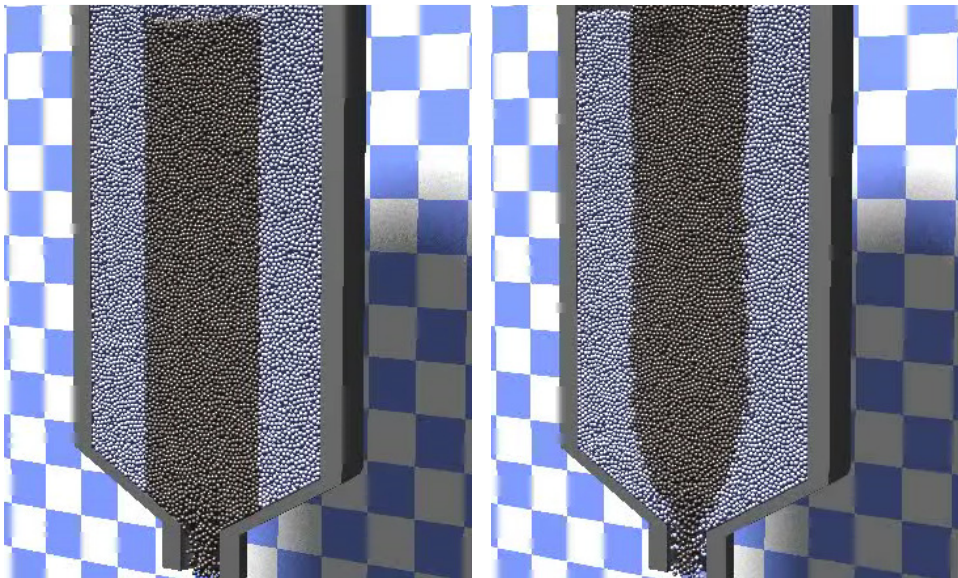


Figura 1: simulazione del *refueling* in reattore PBR (180'000 corpi rigidi con attrito)

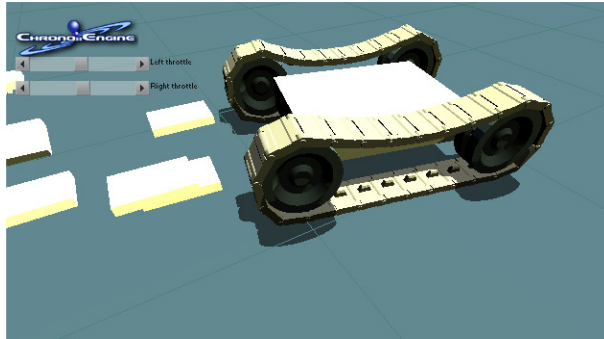


Figura 2: simulazione di veicolo cingolato in real-time

produrre idrogeno in grandi quantità. Il combustibile fissile, lambito da elio a 1000°C, è contenuto in centinaia di migliaia di sfere in lenta ricircolazione all'interno di un silo; il movimento delle sfere produce un flusso granulare multi disperso di difficile simulazione. Grazie al nuovo metodo numerico, concretizzatosi in un software multibody, è stato possibile simulare il flusso granulare su un semplice PC laddove in passato era necessario l'impiego di un supercomputer.

3. COLLABORAZIONE CON UNIVERSITY OF WISCONSIN

Dal 2007 è in corso una collaborazione con il laboratorio SBEL dell'University of Wisconsin-Madison, grazie alla cooperazione del Prof.D.Negrut e dei suoi collaboratori. Tale ricerca ha come obiettivo la simulazione della marcia di veicoli su terreni accidentati, con suoli sabbiosi o cedevoli. Obiettivo di tale programma è la simulazione della dinamica di tutte le particelle che costituiscono il terreno, sotto l'effetto di ruote, cingoli o pale; queste ultime si ritrovano nel caso della simulazione di bulldozer e macchine movimento terra (v. Fig.2). Presentandosi milioni di gradi di libertà, si ottiene un problema di non-smooth-dynamics dalla complessità difficilmente attaccabile con metodi computazionali già noti; per tale ragione si è sviluppato un algoritmo *matrix-less* che può essere eseguito su piattaforme ad elevato parallelismo.

I risultati preliminari di questo studio hanno stimolato l'interesse di importanti società multinazionali operanti nel settore del movimento terra, nonché di altri centri di ricerca USA intenzionati ad usare il nuovo metodo numerico per la simulazione di veicoli-sonda per esplorazione spaziale.

4. COLLABORAZIONE CON ISPE

Da alcuni anni, per interessamento dell'Ing. Marco Silvestri, sono in corso attività di ricerca e didattica in collaborazione con l'International Society of Pharmaceutical Engineering (ISPE). Grazie ai contatti internazionali così avviati, si è proposto un corso post-laurea orientato alla tematica delle Good Automated Manufacturing Practices (GAMP). Conoscenze, processi e normative contenute in tale ambito sono di particolare importanza ai fini della progettazione delle macchine automatiche che operano nei campi sottoposti al controllo della Food and Drug Administration.

Si sono inoltre attivate ricerche in merito al trasferimento dell'approccio GAMP dall'industria farmaceutica all'industria alimentare, con interessanti effetti nell'ambito industriale emiliano, tradizionalmente sede di importanti aziende che sviluppano macchine per imbottigliamento e packaging.

Per quanto concerne gli sviluppi futuri delle nostre attività nel settore dell'ingegneria farmaceutica, si prevede di realizzare un'officina farmaceutica dedicata alla produzione di farmaci particolari, sita all'interno del Parco Scientifico e Tecnologico dell'Università di Parma. Tale struttura coinvolgerà il Dipartimento di Farmacia e permetterà a studenti e tesisti di ingegneria, ad esempio afferenti al curriculum automazione, di svolgere attività di tirocinio e di ricerca.

5. COLLABORAZIONE CON SKF NETHERLANDS

Sulla scorta di una lunga esperienza maturata nell'ambito della tribologia, il gruppo di ricerca di Meccanica Applicata dell'Università di Parma ha recentemente instaurato una collaborazione con la sede olandese della SKF in merito allo sviluppo di sensori per tenute a contatto circolare.

Il sensore, rappresentato in Fig.3, è stato inizialmente concepito per ricerche interne al nostro dipartimento; in seguito è stato ingegnerizzato e prodotto in diverse copie per conto della SKF che ha mostrato interesse per questa soluzione. In particolare, il sensore da noi concepito è in grado di rilevare lo spostamento micrometrico in 2D di un punto di una tenuta a labbro grazie ad un ago supportato da un sottile telaio dotato di estensimetri. La versione perfezionata ed ingegnerizzata di questo sensore è realizzata con processo di elettroerosione; il software è sviluppato su piattaforma LabView ed è stata messa a punto una procedura di calibrazione per la serie realizzata.

6. COLLABORAZIONE CON MC GILL UNIVERSITY

Il nostro dipartimento è dotato di un laboratorio di vibrazioni meccaniche: tale ambito scientifico è da tempo gestito dal Prof. M.Amabili che ha attivato diversi temi di ricerca in collaborazione con la Mc Gill University, in Canada.

Le ricerche vertono principalmente su problemi di vibrazioni di piastre, gusci ed edifici. Ci si avvale di metodi numerici e sperimentali, questi ultimi supportati da avanzate



Figura 3: apparato sperimentale sviluppato per la SKF

strumentazioni quali vibrometri laser, laser doppler, etc. L'attività svolta in tale contesto ha prodotto recenti ricerche su metodi non distruttivi per la valutazione dell'integrità strutturale degli edifici storici.

7. COLLABORAZIONE CON NVIDIA CORPORATION

La tecnologia GP-GPU (General Purpose Graphical Processor Unit) è un recente approccio al calcolo parallelo nel quale si utilizzano i moderni multiprocessori GPU delle schede grafiche, usualmente impiegati per operazioni di *pixel shading*, per svolgere calcoli di tipo scientifico con prestazioni superiori di uno o due ordini di grandezza, in termini di velocità, rispetto alle implementazioni su processori CPU tradizionali. I processori GPU delle schede grafiche più potenti sono in grado di eseguire migliaia di *threads* in parallelo; tuttavia il loro impiego non è immediato e ciò ha motivato lo sviluppo di strumenti, esempi e documentazione da parte dei principali produttori, ovvero Nvidia ed ATI. In particolare l'Nvidia Corporation, con sede in USA, ha prodotto la serie di schede GPU "Tesla" specificamente dedicate al calcolo scientifico; inoltre distribuisce strumenti di sviluppo software, denominati CUDA, per agevolare lo sviluppo di algoritmi paralleli su schede GPU.

La società Nvidia sostiene numerose attività di ricerca universitaria nel mondo attraverso il Professor Partnership Program. Dal 2007 l'Ing.A.Tasora fa parte di tale programma ed ha ricevuto, come sponsorizzazione, numerose schede Tesla e prototipi di schede GPU per una potenza di calcolo superiore a 5 TFlops (equivalente, per confronto, a quattro supercomputer ASCI Red). L'interesse dell'Nvidia per le nostre attività di ricerca deriva dal fatto che l'algoritmo di simulazione multibody già esposto nei paragrafi precedenti si presta ad un'implementazione su architettura parallela di tipo GPU.

Lo sviluppo della versione GPU del software multibody è svolta in collaborazione con l'University of Wisconsin, dato che le prestazioni di supercalcolo permetterebbero di risolvere agevolmente i problemi di interazione ruota-terreno di cui si è già parlato nei paragrafi precedenti. I primi risultati mostrano un incremento di un ordine di grandezza nella velocità del solutore.



Figura 4: una delle otto schede GPU/Tesla per supercalcolo presenti nei nostri laboratori

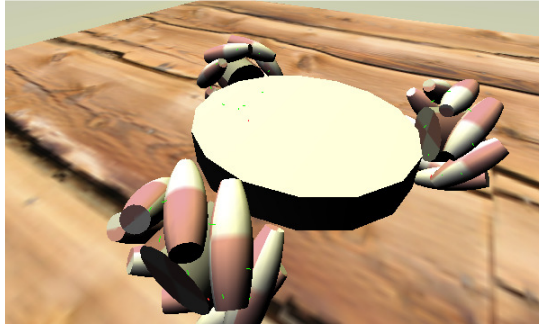


Figura 5: simulazione del robot mobile con ruote omnidirezionali

8. COLLABORAZIONE CON VUT BRNO

Dal 2006 è in corso un rapporto di collaborazione col Dipartimento di Sistemi Robotici dell'Università di Brno (istituto VUT), in Repubblica Ceca. In tale sede sono state svolte diverse attività didattiche -seminari, lezioni- grazie al programma Socrates che supporta lo scambio di docenti in ambito europeo. Inoltre si è iniziata un'attività di ricerca che riguarda lo studio di robot mobili con ruote omnidirezionali (vedi Fig.5); un prototipo è in fase di costruzione presso il VUT e appositi algoritmi di simulazione real-time sono allo studio presso il nostro Dipartimento.

9. CONCLUSIONI

Consapevole della dimensione internazionale della scienza moderna ed incoraggiato dall'esempio delle positive esperienze maturate fino ad ora, il nostro gruppo di ricerca prevede di proseguire ed intensificare i rapporti con università, aziende e laboratori esteri nel prossimo futuro; ciò sarà fatto sia al fine di produrre ulteriore attività di ricerca, sia per innovare ed arricchire la didattica con contenuti innovativi.

BIBLIOGRAFIA

- [1] Tasora, A., Anitescu, M., 2008. "A Fast NCP Solver for Large Rigid-Body Problems with Contacts, Friction, and Joints", *Multibody Dynamics: Computational Methods And Applications, Series: Computational Methods in Applied Sciences*, Bottasso, C.L. Editor, Springer, pp. 45-54.
- [2] Tasora, A., Negrut, D., Anitescu, M., 2008. "Large-Scale Parallel Multibody Dynamics with Frictional Contact on the Graphical Processing Unit", *Journal Of Multi-Body Dynamics*, Vol. 222, No. 4., pp. 315-326.
- [3] Silvestri, M., Prati, E., Tasora, A., 2005. "Numerical and Experimental Study of Dynamic Elastomeric Seals Behaviour under Actively Controlled Temperature Conditions", in *Proceedings of 18th International Conference on Fluid Sealing*, Antwerp, Belgium, p. 1-10.
- [4] Silvestri, M., Prati, E., Tasora, A., 2004. "Dynamic Seals Behaviour under Effect of Radial

Vibration”, in Proceedings of *14th International Colloquium on Tribology and Lubricants Engineering*, pp. 1-8.

[5] Negrut, D., Anitescu, M., Tasora, A., 2008. “Large-Scale Parallel Multibody Dynamics with Frictional Contact on the GPU”, in Proceedings of *ASME Dynamic Systems and Control Conference*. Ann Arbor, Michigan, USA.

[6] Silvestri, M., Prati, E., Tasora, A. 2007. “Frictional Behavior of Radial Lip Seals under Dynamic Operating Conditions for Different Elastomer Rings”, In Proceedings of *19th International Conference on Fluid Sealing - BHR Group*, Poitiers, pp.1-12.

INDICE DEGLI AUTORI

Andrisano Angelo Oreste	153
Ascari Alessandro	193
Baldini Andrea	103
Barbanti Giovanni	153
Benedetti Angela	283
Berselli Giovanni	271
Campana Giampaolo	193
Carricato Marco	1
Castagnetti Davide	57
Cocconcelli Marco	327
Cuccolini Gabriele	193
Dalpiaz Giorgio	13, 293
Delvecchio Simone	293
Di Gregorio Raffaele	221
Donati Lorenzo	165
Dragoni Eugenio	57, 235
D'Elia Gianluca	293
Faretra Marco	153
Fortunato Alessandro	193
Freddi Alessandro	67
Ghelli Daniele	93, 133
Gherardini Francesco	153
Giacopini Matteo	103
Innocenti Carlo	261
Leali Francesco	153
Macchelli Alessandro	121
Maggiore Alberto	313
Mantega Claudio	181
Martini Alberto	1
Medri Gianluca	113
Meneghetti Umberto	313
Minak Giangiacomo	93, 133

Molari Pier Gabriele	283
Morelli Piero	283
Morri Andrea	207
Mucchi Emiliano	13
Olmi Giorgio	67
Orazi Leonardo	193
Parenti Castelli Vincenzo	271
Pellicciari Marcello	153
Pini Fabio	153
Prati Edzeario	337
Rivola Alessandro	1
Rosi Roberto	103
Rubini Riccardo	327
Schikorra Marco	165
Scirè Mammano Giovanni	235
Silvestri Marco	337
Spaggiari Andrea	57
Spinella Igor	235
Squatrito Rosario	207
Strozzi Antonio	103
Tani Giovanni	181, 193
Tasora Alessandro	337
Tekkaya A. Erman	165
Todaro Ivan	207
Tomesani Luca	165, 207
Troncossi Marco	1
Vassura Gabriele	121, 271
Vergnano Alberto	153
Vertechy Rocco	271
Zanarini Alessandro	31
Zucchelli Andrea	93, 133